

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: September 27, 2004, 07:46:34 ; Search time 22285 Seconds

(without alignments)  
11591.855 Million cell updates/sec

Title: US-10-030-294-2

Perfect score: 5960

Sequence: 1 cagcgttcaggatgctcgtc.....tttggctgatatcgcgcgc

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmb1:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_str:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_str:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pin:\*  
35: em\_hcg\_rnd:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_higo\_hum:\*  
40: em\_higo\_mus:\*  
41: em\_higo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5960	100.0	5960	6 BD093090	BD093090 Gene enco
2	5919.6	99.3	152646	6 AC078809	AC078809 Homo sapi
3	5681.6	95.3	13940	6 BD093100	BD093100 Gene enco
4	5676	95.2	5676	6 BD093089	BD093089 Gene enco
5	5676	95.2	5676	6 BD093091	BD093091 Gene enco
6	5676	95.2	5676	6 BD093092	BD093092 Gene enco
7	843.2	14.1	160574	2 AC135489	AC135489 Rattus no
8	843.2	14.1	284005	2 AC113784	AC113784 Rattus no
9	835.8	14.0	204313	2 AC126304	AC126304 Rattus no
10	835.8	14.0	218852	10 AC113020	AC113020 Mus muscu
11	499	8.4	835	9 H816301	H816301 Homo sapi
12	470.4	7.9	531	9 AY339663S1	AY339663 Homo sapi
13	402.2	6.7	56268	2 AC091240	AC091240 Mus muscu
14	273.4	4.6	208158	10 AF533892	AF533892 Mus muscu
15	212.4	3.6	159249	9 AC103882	AC103882 Homo sapi
16	205	3.4	60323	2 AC087334	AC087334 Homo sapi
17	204.4	3.4	173456	9 AL158198	AL158198 Human DNA
18	201.2	3.4	150846	2 AC036173	AC036173 Homo sapi
19	201.6	3.4	178178	2 AC135613	AC135613 Pat trogl
20	200	3.4	195808	2 AL591182	AL591182 Homo sapi
21	196.6	3.3	125377	2 AC118135	AC118135 Homo sapi
22	196.6	3.3	181442	2 AC134875	AC134875 Homo sapi
23	196.6	3.3	202827	9 AC016907	AC016907 Homo sapi
24	196.2	3.3	219122	9 AC080038	AC080038 Homo sapi
25	196	3.3	148911	2 AL353142	AL353142 Homo sapi
26	196	3.3	273403	9 AC011498	AC011498 Homo sapi
27	195	3.3	47028	9 AC117488	AC117488 Homo sapi
28	194	3.3	163246	2 AC068401	AC068401 Homo sapi
29	194	3.3	174428	9 AC005821	AC005821 Homo sapi
30	193.4	3.2	118384	9 AF389851	AF389851 Homo sapi
31	193.2	3.2	196467	9 AC022352	AC022352 Homo sapi
32	193	3.2	139218	2 AC019289	AC019289 Homo sapi
33	193	3.2	151880	9 AP005716	AP005716 Homo sapi
34	193	3.2	166093	9 AC106013	AC106013 Homo sapi
35	193	3.2	175931	2 AC026618	AC026618 Homo sapi
36	193	3.2	180776	2 AC032043	AC032043 Homo sapi
37	193	3.2	183336	9 AC108688	AC108688 Homo sapi
38	193	3.2	229164	2 AC092893	AC092893 Homo sapi
39	191.8	3.2	179305	2 AC144411	AC144411 Homo sapi
40	191.2	3.2	131757	9 AC004540	AC004540 Homo sapi
41	190.6	3.2	78340	9 AL359473	AL359473 Human DNA
42	190.6	3.2	167416	2 AC026857	AC026857 Homo sapi
43	190.6	3.2	218074	2 H0AC02044	AC02044 Human Chr
44	190.4	3.2	157136	2 AC068251	AC068251 Homo sapi
45	190.4	3.2	165240	9 AP002795	AP002795 Homo sapi

## ALIGNMENTS

RESULT 1  
BD093090 5960 bp DNA linear PAT 27-AUG-2002  
LOCUS BD093090  
DEFINITION Gene encoding promoter region of tumor suppressor gene p51.  
ACCESSION BD093090.1 GI:22638678  
VERSION WD0100818-A/2.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 5960)  
AUTHORS Sakai,T., Kagaya,S., Sato,T., Sukenaga,Y. and Fujii,E.  
TITLE Gene encoding promoter region of tumor suppressor gene p51  
JOURNAL Patent: WO 0100818-A 2 04-JAN-2001.

## COMMENT

TOSHIYUKI SAKAI, NIPON KAVAKU KK, SHIGEHIDE KAGAWA, TAKAMICHI SATO,  
YOSHIAZU SUKENAGA, HIDEJI FUJII  
OS Homo sapiens (human)  
FN WO 0100818-A/2  
PD 04-JUN-2001  
PE 28-JUN-2000 WO 2000JP004261  
PR 29-JUN-1999 JP 99P 183195  
PI TOSHIYUKI SAKAI, SHIGEHIDE KAGAWA, TAKAMICHI SATO, YOSHIAZU SAKAI,  
SUENAGA,  
PI HIDEJI FUJII  
PC C12N15/10, C12N5/10, C12N1/68, A61K48/00, A61K45/00, A61P43/00, PC  
A61P35/00

## FEATURES

CC  
FH Key Location/Qualifiers  
FT promoter (1).. (5676)  
FT 5'UTR (5677).. (5960)  
FT Intron (5677).. (5960)  
Location/Qualifiers  
1..5960

## ORIGIN

Query Match 100.0%; Score 5960; DB 6; Length 5960;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5960; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGCTGTTGAGGAGATGCTGAGAAAGAAAGAGCCACCATTCCTTCTGACACTGGGTGT 60  
DB 1 CAGCTGTTGAGGAGATGCTGAGAAAGAAAGAGCCACCATTCCTTCTGACACTGGGTGT 60  
QY 61 GACTTTGAGGAGATGCTGAGAAAGAAAGAGCCACCATTCCTTCTGACACTGGGTGT 120  
DB 61 GACTTTGAGGAGATGCTGAGAAAGAAAGAGCCACCATTCCTTCTGACACTGGGTGT 120  
QY 121 TGGCTCTGCTCCCTGACATGCTCTTCTCTGAGACATCTCCTTAAAGCATCTTCTG 180  
DB 121 TGGCTCTGCTCCCTGACATGCTCTTCTCTGAGACATCTCCTTAAAGCATCTTCTG 180  
QY 181 ACATTAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240  
DB 181 ACATTAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240  
QY 241 GTCAATTCAGAAAG 300  
DB 241 GTCAATTCAGAAAG 300  
QY 301 AGTCATTGCAATACCAATTCAGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 360  
DB 301 AGTCATTGCAATACCAATTCAGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 360  
QY 361 GTCAAGCTGAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420  
DB 361 GTCAAGCTGAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420  
QY 421 CAGAACTTGGCTTCCCTTAACTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480  
DB 421 CAGAACTTGGCTTCCCTTAACTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480  
QY 481 CAGAAAGTGTCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540  
DB 481 CAGAAAGTGTCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540  
QY 541 TGGTGAATTTACTGACACACCT 600  
DB 541 TGGTGAATTTACTGACACACCT 600  
QY 601 AATGCAAGATGACTGAGAAAGCTAGAGCTGAGGTTCAAGTTACCTGACACAAAGAAAT 660  
DB 601 AATGCAAGATGACTGAGAAAGCTAGAGCTGAGGTTCAAGTTACCTGACACAAAGAAAT 660  
QY 661 TCAATTTCTGATCTCATAGTCAAGGCTGACAGAGCTTACGGAACATGCAAGATCAT 720

DB 661 TCAATTTCTGATCTCATAGTCAAGGCTGACAGAGCTTACGGAACATGCAAGATCAT 720  
QY 721 CTGCTTAAAGCTTCTGAGAGATCTGTTCTTCTGACATGCTGACATGCTGACATGCTGAC 780  
DB 721 CTGCTTAAAGCTTCTGAGAGATCTGTTCTTCTGACATGCTGACATGCTGACATGCTGAC 780  
QY 781 TTCCTGTTTAAAGAACTTTATTTCTTGAATACTCTGCTGACATGCTGACATGCTGAC 840  
DB 781 TTCCTGTTTAAAGAACTTTATTTCTTGAATACTCTGCTGACATGCTGACATGCTGAC 840  
QY 841 CCATGATTCATGATCAGGCTCTCTCTGACAAACATGCAATCTTCTTGGGAAT 900  
DB 841 CCATGATTCATGATCAGGCTCTCTCTGACAAACATGCAATCTTCTTGGGAAT 900  
QY 901 TGAATCTTAAGCTGAATAGCTGAAGTTCAAAAAGCTTGAATCTGAATCTGAATCTGA 960  
DB 901 TGAATCTTAAGCTGAATAGCTGAAGTTCAAAAAGCTTGAATCTGAATCTGAATCTGA 960  
QY 961 AGTGGCTTGGAAAG 1020  
DB 961 AGTGGCTTGGAAAG 1020  
QY 1021 CCTGGTTACAGCCCTTCTGAGAGTGTGTTTCTTCAAGCTGCTTAAATAGTGT 1080  
DB 1021 CCTGGTTACAGCCCTTCTGAGAGTGTGTTTCTTCAAGCTGCTTAAATAGTGT 1080  
QY 1081 GAATTTCAATTTCTTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 1140  
DB 1081 GAATTTCAATTTCTTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 1140  
QY 1141 GCTTTAGTGAAG 1200  
DB 1141 GCTTTAGTGAAG 1200  
QY 1201 TCTTCCCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260  
DB 1201 TCTTCCCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260  
QY 1261 TCTGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1320  
DB 1261 TCTGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1320  
QY 1321 AACATTTCTCTCTCTTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380  
DB 1321 AACATTTCTCTCTCTTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380  
QY 1381 CAAAGACCTCAG 1440  
DB 1381 CAAAGACCTCAG 1440  
QY 1441 AGTTTGTGACCTTCTGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500  
DB 1441 AGTTTGTGACCTTCTGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500  
QY 1501 AAACAAACAAACAGAGATTCACAGGTTCAAGCTTAAATTAATTAATTAATTAATTAAT 1560  
DB 1501 AAACAAACAAACAGAGATTCACAGGTTCAAGCTTAAATTAATTAATTAATTAATTAAT 1560  
QY 1561 AGTTGACACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1620  
DB 1561 AGTTGACACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1620  
QY 1621 ACGAAGAAATGAG 1680  
DB 1621 ACGAAGAAATGAG 1680  
QY 1681 CATAGTACAGAGTAAAGCTCAATGCTTAAATTAATTAATTAATTAATTAATTAATTAAT 1740  
DB 1681 CATAGTACAGAGTAAAGCTCAATGCTTAAATTAATTAATTAATTAATTAATTAATTAAT 1740  
QY 1741 ATAAAAATTAAGCTTCAATGAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1800

**BEST AVAILABLE COPY**

[illegible]

QY	2881	CTTGCAACATCAAAAGTAAATATTCATAGCATTTGTTATGTAACCTTAGTCCCTCAAGAAAT	2946
Db	2881	CTTGCAACATCAAAAGTAAATATTCATAGCATTTGTTATGTAACCTTAGTCCCTCAAGAAAT	2946
QY	2941	AAGATGAACATAAATAATTAAGAAATTAATATGTAATTAATATATGATGCTCACTTAT	3000
Db	2941	AAGATGAACATAAATAATTAAGAAATTAATATGTAATTAATATATGATGCTCACTTAT	3000
QY	3001	TTTCACTGATGCTGATAGTTGATGCACTGATGTTCTGTGTAATTTCAAGAGTTG	3060
Db	3001	TTTCACTGATGCTGATAGTTGATGCACTGATGTTCTGTGTAATTTCAAGAGTTG	3060
QY	3061	CCCTCTTTTCCGGGTCAACTCTGGCAATTTATTCATATGCAATAGAGCCAACTCT	3120
Db	3061	CCCTCTTTTCCGGGTCAACTCTGGCAATTTATTCATATGCAATAGAGCCAACTCT	3120
QY	3121	TTTTCAATATTACTTAATTTAAATTTGTATGCAATTAATTTCTGTTCTTAGTCTATTA	3180
Db	3121	TTTTCAATATTACTTAATTTAAATTTGTATGCAATTAATTTCTGTTCTTAGTCTATTA	3180
QY	3181	ACTTGAATTTTTTAAATACACATTTGAATATGACATATGATCTTAATATTAAT	3240
Db	3181	ACTTGAATTTTTTAAATACACATTTGAATATGACATATGATCTTAATATTAAT	3240
QY	3241	TAAATACGTTAGGCTATATAACCTTTTAAATTTTTTAAATAAATAGATGATGCTGCT	3300
Db	3241	TAAATACGTTAGGCTATATAACCTTTTAAATTTTTTAAATAAATAGATGATGCTGCT	3300
QY	3301	CATGCTGTATATCCCAACACTTGGGAAGCCGGTGGGAGCATCTTAGTATCCAGAG	3360
Db	3301	CATGCTGTATATCCCAACACTTGGGAAGCCGGTGGGAGCATCTTAGTATCCAGAG	3360
QY	3361	TTTGAGACCAATGACGGGCAACACAGCAAGACCCATATCTTAATAAACAATAAACA	3420
Db	3361	TTTGAGACCAATGACGGGCAACACAGCAAGACCCATATCTTAATAAACAATAAACA	3420
QY	3421	AATTAAGTGGATAGTATGTATGTCTACCTGTATGTCAGAGTACACAGAAAGCTGAGGAGA	3480
Db	3421	AATTAAGTGGATAGTATGTATGTCTACCTGTATGTCAGAGTACACAGAAAGCTGAGGAGA	3480
QY	3481	AGCATCACTTGACGCCACGAGAGGTTAGAGGCTGCAGTATCCATATAGCGCTGCTACCTC	3540
Db	3481	AGCATCACTTGACGCCACGAGAGGTTAGAGGCTGCAGTATCCATATAGCGCTGCTACCTC	3540
QY	3541	AGTCTGGGTGACAGTGCACAAAGCGTCTCAAAAATTAATTAATTAATAAATTAATTAATTA	3600
Db	3541	AGTCTGGGTGACAGTGCACAAAGCGTCTCAAAAATTAATTAATTAATAAATTAATTAATTA	3600
QY	3601	AAAAACAAAAATTAATTAATTTTAAAAACACACACTATAGAGATGTTGCAATTA	3660
Db	3601	AAAAACAAAAATTAATTAATTTTAAAAACACACACTATAGAGATGTTGCAATTA	3660
QY	3661	TTATTTGGAGTCTATATCCCTGGAAGTTAAATTTAAATATTTTGAAGAGTTCTTCTCA	3720
Db	3661	TTATTTGGAGTCTATATCCCTGGAAGTTAAATTTAAATATTTTGAAGAGTTCTTCTCA	3720
QY	3721	TTTCTTGAAGAGCTGCATTTGTAATATCAAGCTATAGAGAAACACTAGGGCTCGCACT	3780
Db	3721	TTTCTTGAAGAGCTGCATTTGTAATATCAAGCTATAGAGAAACACTAGGGCTCGCACT	3780
QY	3781	CCAAAGTGTGTCAGAGACACAGACAGATCAAGTATCTGGGAAGCTGTTAGAAATGACG	3840
Db	3781	CCAAAGTGTGTCAGAGACACAGACAGATCAAGTATCTGGGAAGCTGTTAGAAATGACG	3840
QY	3841	AGCTTTAGGCTTACATCCCAAGACCTTACTGACACAAATCTGATTAACAAATTTCTAGCT	3900
Db	3841	AGCTTTAGGCTTACATCCCAAGACCTTACTGACACAAATCTGATTAACAAATTTCTAGCT	3900
QY	3901	GCTCAAGGGCATTAAATCTTGAAGAGCTTGCATAGAAATCTTCACTCACTTTC	3960
Db	3901	GCTCAAGGGCATTAAATCTTGAAGAGCTTGCATAGAAATCTTCACTCACTTTC	3960

Thu Sep 30 11:51:28 2004

us-10-030-294-2.Std.rge

Page 4

QY 3961 ATTATTAATGGAATCACTTGCGCTGCTGACAGGAATGATTTTATTTATTCAGAA 4020  
| | | | |  
DB 3961 ATTATTAATGGAATCACTTGCGCTGCTGACAGGAATGATTTTATTTATTCAGAA 4020  
QY 4021 CCTTCTATTTAGGTCATCTAATTTGCTAAATGCGAGGAAGAAAGCCAACTCTTAACT 4080  
| | | | |  
DB 4021 CCTTCTATTTAGGTCATCTAATTTGCTAAATGCGAGGAAGAAAGCCAACTCTTAACT 4080  
QY 4081 GCAATTTACAAATCTAATTAATTAATGATTAGCAATCTTCCCTTTAACTTCAATTTG 4140  
| | | | |  
DB 4081 GCAATTTACAAATCTAATTAATTAATGATTAGCAATCTTCCCTTTAACTTCAATTTG 4140  
QY 4141 TCGACAGAGCTGTTGATTTGCTGCGGCTGAGCGCGGCTGTTGGAATTTCAAAAT 4200  
| | | | |  
DB 4141 TCGACAGAGCTGTTGATTTGCTGCGGCTGAGCGCGGCTGTTGGAATTTCAAAAT 4200  
QY 4201 CACAGATGTAAGCGGCTCTCGGCTAAGTAAGGAAGAAATGTCAGATTAAATAGCT 4260  
| | | | |  
DB 4201 CACAGATGTAAGCGGCTCTCGGCTAAGTAAGGAAGAAATGTCAGATTAAATAGCT 4260  
QY 4261 TCTCCCTTCATCTGCGCTGAGCAACAATTAATTTTATGAAACATTTTGAGT 4320  
| | | | |  
DB 4261 TCTCCCTTCATCTGCGCTGAGCAACAATTAATTTTATGAAACATTTTGAGT 4320  
QY 4321 TAGATTTACTTAAGGGAATGTCATTTCTGGAAGGCTTAAATGTCACAC 4380  
| | | | |  
DB 4321 TAGATTTACTTAAGGGAATGTCATTTCTGGAAGGCTTAAATGTCACAC 4380  
QY 4381 TTGACATCTAAGTACCTAATTAATTAAGTGTCTGTGACTAGAGGCTGAAGGA 4440  
| | | | |  
DB 4381 TTGACATCTAAGTACCTAATTAATTAAGTGTCTGTGACTAGAGGCTGAAGGA 4440  
QY 4441 AGATGTAAGTCAACGATGTAAGTACGCTTAAATCAAGAGTCTTTTTCCTCCG 4500  
| | | | |  
DB 4441 AGATGTAAGTCAACGATGTAAGTACGCTTAAATCAAGAGTCTTTTTCCTCCG 4500  
QY 4501 TTGAGCTATCTTAAGTACGCTTAAATCAAGAGTCTTTTTCCTCCG 4560  
| | | | |  
DB 4501 TTGAGCTATCTTAAGTACGCTTAAATCAAGAGTCTTTTTCCTCCG 4560  
QY 4561 ACAGGATTAAGTACGCTTAAATCAAGAGTCTTTTTCCTCCG 4620  
| | | | |  
DB 4561 ACAGGATTAAGTACGCTTAAATCAAGAGTCTTTTTCCTCCG 4620  
QY 4621 TCGATTAATTAAGTACGCTTAAATCAAGAGTCTTTTTCCTCCG 4680  
| | | | |  
DB 4621 TCGATTAATTAAGTACGCTTAAATCAAGAGTCTTTTTCCTCCG 4680  
QY 4681 GTCAAGTCAAGCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAA 4740  
| | | | |  
DB 4681 GTCAAGTCAAGCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAA 4740  
QY 4741 CCGAGTCAAGCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAA 4800  
| | | | |  
DB 4741 CCGAGTCAAGCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAA 4800  
QY 4801 GTGACCTCAAGCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAA 4860  
| | | | |  
DB 4801 GTGACCTCAAGCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAA 4860  
QY 4861 TTACTATTAATTAAGTACGCTTAAATCAAGAGTCTTTTTCCTCCG 4920  
| | | | |  
DB 4861 TTACTATTAATTAAGTACGCTTAAATCAAGAGTCTTTTTCCTCCG 4920  
QY 4921 AGGTTATTTTCTTTTATTTGCGGCTGAGGAAAGCTTCAAGGGAATGGAAGGATATCT 4980  
| | | | |  
DB 4921 AGGTTATTTTCTTTTATTTGCGGCTGAGGAAAGCTTCAAGGGAATGGAAGGATATCT 4980  
QY 4981 CTTTCTCTTAAGTCAAGCAAGGAGTGTCTTAAGTTAAATTAATCAAGAAATTTCCCT 5040  
| | | | |  
DB 4981 CTTTCTCTTAAGTCAAGCAAGGAGTGTCTTAAGTTAAATTAATCAAGAAATTTCCCT 5040  
QY 5041 GTCTTTGCTATTTGAGATGTGTACCAACAAGCGGCTTGCTGAAGCAAACTGAAGG 5100  
| | | | |

DB 5041 GTCTTTGCTATTTGAGATGTGTACCAACAAGCGGCTTGCTGAAGCAAACTGAAGG 5100  
QY 5101 CCGGAGGAGGAGGAATGATGAAAAAACAACCAAAACCAAAACCAAAACCAAAACCAAA 5160  
| | | | |  
DB 5101 CCGGAGGAGGAGGAATGATGAAAAAACAACCAAAACCAAAACCAAAACCAAAACCAAA 5160  
QY 5161 ACAAAACATTTTACCCCAAGAAATAGTCAAGAAATCTCAATTAATCAAGATTCAGA 5220  
| | | | |  
DB 5161 ACAAAACATTTTACCCCAAGAAATAGTCAAGAAATCTCAATTAATCAAGATTCAGA 5220  
QY 5221 TACAGAGAGGTTATGTAAGTCAAGAGGAGTGTGACATCATAGCTTAATTCAGTTACA 5280  
| | | | |  
DB 5221 TACAGAGAGGTTATGTAAGTCAAGAGGAGTGTGACATCATAGCTTAATTCAGTTACA 5280  
QY 5281 AAGTCCAGGCTGCAATTAATTAATCTGATGCAATTAATGCAAGATCAATCAAGACA 5340  
| | | | |  
DB 5281 AAGTCCAGGCTGCAATTAATTAATCTGATGCAATTAATGCAAGATCAATCAAGACA 5340  
QY 5341 GAGATCAAGATTAAGAGATGCTCCAGGCTCCAAATTTGCAAAACCAAGTGTGCTACTA 5400  
| | | | |  
DB 5341 GAGATCAAGATTAAGAGATGCTCCAGGCTCCAAATTTGCAAAACCAAGTGTGCTACTA 5400  
QY 5401 TACGTCAGAGCTTGAAGCGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5460  
| | | | |  
DB 5401 TACGTCAGAGCTTGAAGCGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5460  
QY 5461 GCTGTCAGAGCTTGAAGCGCTTGAAGAGTGTGATGATGATGATGATGATGATGATGAT 5520  
| | | | |  
DB 5461 GCTGTCAGAGCTTGAAGCGCTTGAAGAGTGTGATGATGATGATGATGATGATGATGAT 5520  
QY 5521 TGAAGAAAGCTCCGCTCTTTCGAAATTAATTAATGAGAGAGAGAGAGAGAGAGAGAG 5580  
| | | | |  
DB 5521 TGAAGAAAGCTCCGCTCTTTCGAAATTAATTAATGAGAGAGAGAGAGAGAGAGAGAG 5580  
QY 5581 TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5640  
| | | | |  
DB 5581 TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5640  
QY 5641 AGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5700  
| | | | |  
DB 5641 AGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5700  
QY 5701 GGAATTAATTTGAACTTCAAGGAGTGTGCAAGCTTCAAGGAGTGTGCAAGGAGTGTG 5760  
| | | | |  
DB 5701 GGAATTAATTTGAACTTCAAGGAGTGTGCAAGCTTCAAGGAGTGTGCAAGGAGTGTG 5760  
QY 5761 CCGAGGAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGTGTG 5820  
| | | | |  
DB 5761 CCGAGGAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGTGTG 5820  
QY 5821 ATTATTAATTTGAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGT 5880  
| | | | |  
DB 5821 ATTATTAATTTGAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGT 5880  
QY 5881 TATTAATTTGAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGT 5940  
| | | | |  
DB 5881 TATTAATTTGAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGT 5940  
QY 5941 TTTTGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5960  
| | | | |  
DB 5941 TTTTGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5960

RESULT 2  
AC07809/c 162646 bp DNA linear PRI 25-JAN-2003  
LOCUS AC07809 Homo sapiens 3 BAC Rpl1-60063 (Rosewell Park Cancer Institute Human  
DEFINITION BAC Library) complete sequence.  
ACCESSION AC07809  
VERSION AC07809.9 GI:1903390  
KEYWORDS HTG  
SOURCE Homo sapiens (human)

BEST AVAILABLE COPY

## ORGANISM

## REFERENCE

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 162646)

## TITLE

Direct Submission  
Submitted (25-JAN-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Mar 1, 2002 this sequence version replaced gi:117977513.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
gc-help@bcm.tmc.edu

## COMMENT

CLONE LENGTH: This sequence does not necessarily represent the  
entire insert of this clone. Overlapping regions of clones are only  
sequenced and submitted once, so the sequence for the remainder of  
the insert may be found in the record for the adjacent clones.  
Overlapping clones are noted at the beginning and end of the  
Features listing.

## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches  
of a local database that includes entries from dbSTS, GDB, and  
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,  
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST  
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-14) to the  
EST and cDNA sequences. Genes demonstrate at least two exons  
flanked by consensus splice sites that maintained sequence  
continuity across the splice junctions. Sequences that are not  
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
standard of double strand coverage with a minimum of 2 clones and 2  
reads with no ambiguities or 2 chemistries with a minimum of 2  
clones and 3 reads with no ambiguities. If the sequence quality for  
a region does not meet this standard, it will be indicated in the  
annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
standards - estimated error rate less than 1 per 10,000 bases.  
Reports of lowest quality individual bases and measures of base  
quality are listed below. Description of the metrics can be found  
at URL:  
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.htm>.

## QUALSTAT-REPORT.

## FEATURES

## source

1 162646  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-600G3"

## misc\_feature

1..2004  
/note="overlaps bases 79511..81510 of clone AC117486"  
/function="clone overlap"  
complement(248..339)  
/rpt\_family="MIR"

## repeat\_region

491..556  
/rpt\_family="TNA"  
/rpt\_family="TNA"  
complement(84..961)  
/rpt\_family="TNA"

## repeat\_region

1586..1727  
/rpt\_family="MIR"  
/rpt\_family="MIR"

## repeat\_region

1683..1736  
/rpt\_family="T2"  
/rpt\_family="T2"

## repeat\_region

2734..2922  
/rpt\_family="MIR"  
/rpt\_family="MIR"

## repeat\_region

2977..3083  
complement(1383..3562)  
/rpt\_family="MIR"  
/rpt\_family="MIR"

## repeat\_region

3742..7012  
/rpt\_family="L1MC3"  
/rpt\_family="L1MC3"

## repeat\_region

7013..7275  
/rpt\_family="AluSx"

## ORGANISM

## REFERENCE

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 162646)

## TITLE

Direct Submission  
Submitted (25-JAN-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Mar 1, 2002 this sequence version replaced gi:117977513.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
gc-help@bcm.tmc.edu

## COMMENT

CLONE LENGTH: This sequence does not necessarily represent the  
entire insert of this clone. Overlapping regions of clones are only  
sequenced and submitted once, so the sequence for the remainder of  
the insert may be found in the record for the adjacent clones.  
Overlapping clones are noted at the beginning and end of the  
Features listing.

## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches  
of a local database that includes entries from dbSTS, GDB, and  
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,  
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST  
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-14) to the  
EST and cDNA sequences. Genes demonstrate at least two exons  
flanked by consensus splice sites that maintained sequence  
continuity across the splice junctions. Sequences that are not  
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
standard of double strand coverage with a minimum of 2 clones and 2  
reads with no ambiguities or 2 chemistries with a minimum of 2  
clones and 3 reads with no ambiguities. If the sequence quality for  
a region does not meet this standard, it will be indicated in the  
annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
standards - estimated error rate less than 1 per 10,000 bases.  
Reports of lowest quality individual bases and measures of base  
quality are listed below. Description of the metrics can be found  
at URL:  
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.htm>.

## QUALSTAT-REPORT.

## FEATURES

## source

1 162646  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-600G3"

## misc\_feature

1..2004  
/note="overlaps bases 79511..81510 of clone AC117486"  
/function="clone overlap"  
complement(248..339)  
/rpt\_family="MIR"

## repeat\_region

491..556  
/rpt\_family="TNA"  
/rpt\_family="TNA"  
complement(84..961)  
/rpt\_family="TNA"

## repeat\_region

1586..1727  
/rpt\_family="MIR"  
/rpt\_family="MIR"

## repeat\_region

1683..1736  
/rpt\_family="T2"  
/rpt\_family="T2"

## repeat\_region

2734..2922  
/rpt\_family="MIR"  
/rpt\_family="MIR"

## repeat\_region

2977..3083  
complement(1383..3562)  
/rpt\_family="MIR"  
/rpt\_family="MIR"

## repeat\_region

3742..7012  
/rpt\_family="L1MC3"  
/rpt\_family="L1MC3"

## repeat\_region

7013..7275  
/rpt\_family="AluSx"

Thu Sep 30 11:51:28 2004

us-10-030-294-2.Bld.rge

Page 6

```
repeat_region 7276..7892
/rpc_family="L1MC3"
complement(8145..8365)
repeat_region 9447..9479
/rpc_family="MIR"
repeat_region 9648..9780
/rpc_family="AT_rich"
repeat_region 10440..10660
/rpc_family="MER20"
complement(10703..11014)
repeat_region 11431..11683
/rpc_family="AluSx"
/rpc_family="AluSx"
standard name="79189"
complement(11911..12013)
repeat_region 12654..12964
/rpc_family="MIR"
repeat_region 13022..13142
/rpc_family="AluSx"
repeat_region 15041..15337
/rpc_family="MIR"
repeat_region 15456..15480
/rpc_family="AluSg"
complement(16012..16151)
repeat_region 17138..17449
/rpc_family="MIR"
repeat_region 17765..17895
/rpc_family="AluSx"
repeat_region 18694..18860
/rpc_family="L2"
repeat_region 18907..19141
/rpc_family="MIR"
repeat_region 19474..19515
/rpc_family="MIR"
repeat_region 20726..21393
/rpc_family="TA)n"
complement(21402..20118)
repeat_region 21710..23051
/rpc_family="L1"
repeat_region 23048..24606
/rpc_family="L1M1"
repeat_region 24611..24765
/rpc_family="L1M1"

Query Match 99.3% Score 5919.6; DB 9; Length 162646;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 5950; Conservative 0; Mismatches 9; Indels 5; Gaps 2;
```

```
Db 75960 AGCTATTGAAATACATTCATTCAGTTTACTTAAGTTCTAGGCGCCAGCTTACTTCTCATTCAT 75901
Qy 361 GTGACACTGTAGCAAAATATATAGGTCCAAAGTTGGAAGATTAAGCATTCCTTCATGA 420
Db 75900 GTGACACTGTAGCAAAATATATAGGTCCAAAGTTGGAAGATTAAGCATTCCTTCATGA 75841
Qy 421 CAGAACTTGGCTTCCACTTATTAATAATAGATTTGTGTGTGTAGAGTGCAGCATGTA 480
Db 75840 CAGAACTTGGCTTCCACTTATTAATAATAGATTTGTGTGTGTAGAGTGCAGCATGTA 75781
Qy 481 CAGAAATGTGCATTAATTAATAATAGCAAAATCTAGCTCTTCAATTTAGCAATAGT 540
Db 75780 CAGAAATGTGCATTAATTAATAATAGCAAAATCTAGCTCTTCAATTTAGCAATAGT 75721
Qy 541 TGGTAATTTATCTGACACCTCTCTCTTGAAGGTTCTTCTGCTCTCTCATATA 600
Db 75720 TGGTAATTTATCTGACACCTCTCTCTTGAAGGTTCTTCTGCTCTCTCATATA 75661
Qy 601 AATGCAAGATGACCTGGAAGAGCTAGAGACCTGAGTTGATACCTGACCAAGGAAT 660
Db 75660 AATGCAAGATGACCTGGAAGAGCTAGAGACCTGAGTTGATACCTGACCAAGGAAT 75601
Qy 661 TCAATTTCTCTGATCTCATAGTCAAGGCTGCAAGCTCTAGCAATCAAGATCAT 720
Db 75600 TCAATTTCTCTGATCTCATAGTCAAGGCTGCAAGCTCTAGCAATCAAGATCAT 75541
Qy 721 CTGCTTAAAGCTCTGTGTGTGTGCATCTGTGTTCATGACCTGATACCTATTGCTCT 780
Db 75540 CTGCTTAAAGCTCTGTGTGTGTGCATCTGTGTTCATGACCTGATACCTATTGCTCT 75481
Qy 781 TTCCTTGTTAAGAACTTTTATTTTCTTGAAACCTCTGCTGAGTCATGCTTAGG 840
Db 75480 TTCCTTGTTAAGAACTTTTATTTTCTTGAAACCTCTGCTGAGTCATGCTTAGG 75421
Qy 841 CCATGATGCAATGATGAGGCTCTCTGCGCAAAATGAGCATCTTTTGGGAAT 900
Db 75420 CCATGATGCAATGATGAGGCTCTCTGCGCAAAATGAGCATCTTTTGGGAAT 75361
Qy 901 TGAATCTTAAGCTAATAGCTGAAGTTCAAAAAGCTGTGAATCTGAAGCTTAC 960
Db 75360 TGAATCTTAAGCTAATAGCTGAAGTTCAAAAAGCTGTGAATCTGAAGCTTAC 75301
Qy 961 AGTGCTTTGGAAGTACCTGATTCATTTCTTAAGTCCCTGAATTTAATTTAT 1020
Db 75300 AGTGCTTTGGAAGTACCTGATTCATTTCTTAAGTCCCTGAATTTAATTTAT 75241
Qy 1021 CTTGTTAGAGCTTTTCTGAGATGTGTTTTTTTCAAGTGTCTTAAGTCTGT 1080
Db 75240 CTTGTTAGAGCTTTTCTGAGATGTGTTTTTTTCAAGTGTCTTAAGTCTGT 75181
Qy 1081 GAATTTCAATTTCTTTTCAATTCATCT----TTGTTTGTGTGTGTGT 1136
Db 75180 GAATTTCAATTTCTTTTCAATTCATCTTTTGTGTGTGTGTGTGTGTGT 75121
Qy 1137 TTTTGGCTTGAAGTACCAATTCATCTTGTGTTTATACCAAGATTCCTGATTA 1196
Db 75120 TTTTGGCTTGAAGTACCAATTCATCTTGTGTTTATACCAAGATTCCTGATTA 75061
Qy 1197 TAGATCTTCCCTTTAAATTAAGTATCTAAGGCTCAAAAGAGTAAAGCTAGT 1256
Db 75060 TAGATCTTCCCTTTAAATTAAGTATCTAAGGCTCAAAAGAGTAAAGCTAGT 75001
Qy 1257 GAGGCTGAGAGTAAAGTATACAGAGCTGTACTAAGCCAGTTAGCACTGTT 1316
Db 75000 GAGGCTGAGAGTAAAGTATACAGAGCTGTACTAAGCCAGTTAGCACTGTT 74941
Qy 1317 ACACAATTTGCTCTCTCTCTGAGTATAGCAAGTCTTGCAGAAAGAGTACTAT 1376
Db 74940 ACACAATTTGCTCTCTCTCTGAGTATAGCAAGTCTTGCAGAAAGAGTACTAT 74881
Qy 1377 TTGCCAAGACTTGAAGAGCAAGAACTTGTGAGATATGTATGAATCTT 1436
Db 1377 TTGCCAAGACTTGAAGAGCAAGAACTTGTGAGATATGTATGAATCTT 1436
```

Db 74880 TTGCCAAGACCTCAGAGGACCAAGAACAGTTCTGGATATATGATGATGAACCTT 74821  
Qy 1437 AAAAGTTTGTGACCTCTGGCCATATATGTGTATCTAAGACGATTTCTTAT 1496  
Db 74820 AAAAGTTTGTGACCTCTGGCCATATATGTGTATCTAAGACGATTTCTTAT 74761  
Qy 1497 AGCTAAACAAACAAAGAGATCCAGAGTTGAGAGCTATATAGAGTAACTAG 1556  
Db 74760 AGCTAAACAAACAAAGAGATCCAGAGTTGAGAGCTATATAGAGTAACTAG 74701  
Qy 1557 ATACGTTGACACATGATATATCTGAAACCATGAGCTCAAGAGAGAAAAATCC 1616  
Db 74700 ATACGTTGACACATGATATATCTGAAACCATGAGCTCAAGAGAGAAAAATCC 74641  
Qy 1617 AGACAGAGAAATAGTACCAATATGCTGCTATATGATATCTAGCAATGATGT 1676  
Db 74640 AGACAGAGAAATAGTACCAATATGCTGCTATATGATATCTAGCAATGATGT 74581  
Qy 1677 CCACATGATAGCAGTACCTCAATGCTATCAATTAAGTATTAATTAAGT 1736  
Db 74580 CCACATGATAGCAGTACCTCAATGCTATCAATTAAGTATTAATTAAGT 74521  
Qy 1737 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1796  
Db 74520 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 74461  
Qy 1797 AGACGTTGACAAATATGATATATCTTATATACAGAAAGTTCTATGAAAAATCT 1856  
Db 74460 AGACGTTGACAAATATGATATATCTTATATACAGAAAGTTCTATGAAAAATCT 74401  
Qy 1857 CTAGAAATATATACATATCTATATTAACAAAGAGAGTCAATGATGCTAAGGCA 1916  
Db 74400 CTAGAAATATATACATATCTATATTAACAAAGAGAGTCAATGATGCTAAGGCA 74341  
Qy 1917 GGGCTAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1976  
Db 74340 GGGCTAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 74281  
Qy 1977 TGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2036  
Db 74280 TGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 74221  
Qy 2037 CACTGGAATTTGGCAATTTCAATGATGATGATGATGATGATGATGATGATG 2096  
Db 74220 CACTGGAATTTGGCAATTTCAATGATGATGATGATGATGATGATGATG 74161  
Qy 2097 GTCAAGTTTGTGGAATTAATTTTATGATGATGATGATGATGATGATGATG 2156  
Db 74160 GTCAAGTTTGTGGAATTAATTTTATGATGATGATGATGATGATGATG 74101  
Qy 2157 TGCATATACATTAACATGATGATGATGATGATGATGATGATGATGATGATG 2216  
Db 74100 TGCATATACATTAACATGATGATGATGATGATGATGATGATGATGATGATG 74041  
Qy 2217 ACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2276  
Db 74040 ACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 73981  
Qy 2277 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2336  
Db 73980 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 73921  
Qy 2337 AAAAGTGCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2396  
Db 73920 AAAAGTGCATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 73861  
Qy 2397 AAAAGAGAGAAATTTTGTGCTTGAAGATGAGCAATGATGATGATGATG 2456  
Db 73860 AAAAGAGAGAAATTTTGTGCTTGAAGATGAGCAATGATGATGATGATG 73801  
Qy 2457 GTTTTAAACATTTCTGCTGAGAGGAGAGATGATGATGATGATGATGATG 2516  
Db 73800 GTTTTAAACATTTCTGCTGAGAGGAGAGATGATGATGATGATGATGATG 73741

Qy 2517 AATATGCAACCTAGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 2576  
Db 73740 AATATGCAACCTAGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 73681  
Qy 2577 GTAATGAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 2636  
Db 73680 GTAATGAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 73621  
Qy 2637 CCTAGAGTCAAGAAACAAAGATGCTCTTCTCTAGTATGCTAGTGTCCAGT 2696  
Db 73620 CCTAGAGTCAAGAAACAAAGATGCTCTTCTCTAGTATGCTAGTGTCCAGT 73561  
Qy 2697 CCAAAACCTTTCTCTCTTGAATGATGATGATGATGATGATGATGATGATGAT 2756  
Db 73560 CCAAAACCTTTCTCTCTTGAATGATGATGATGATGATGATGATGATGATGAT 73501  
Qy 2757 CACAAACATGATTAACAGGAGGATGATGATGATGATGATGATGATGATGATG 2816  
Db 73500 CACAAACATGATTAACAGGAGGATGATGATGATGATGATGATGATGATGATG 73441  
Qy 2817 CTATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2876  
Db 73440 CTATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 73381  
Qy 2877 CCACCTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2936  
Db 73380 CCACCTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 73321  
Qy 2937 GAATAGAGAGAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2996  
Db 73320 GAATAGAGAGAACTAATTAATTAATTAATTAATTAATTAATTAATTAAT 73261  
Qy 2997 TATTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3056  
Db 73260 TATTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATG 73201  
Qy 3057 TTGCTCTTTTCTGAGTGAATGATGATGATGATGATGATGATGATGATGATG 3116  
Db 73200 TTGCTCTTTTCTGAGTGAATGATGATGATGATGATGATGATGATGATGATG 73141  
Qy 3117 TCTTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3176  
Db 73140 TCTTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 73081  
Qy 3177 AGTAACCTAGAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3236  
Db 73080 AGTAACCTAGAGATTTTAAATTAATTAATTAATTAATTAATTAATTAAT 73021  
Qy 3237 TATTTAAATAGCTTAAAGCTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 3296  
Db 73020 TATTTAAATAGCTTAAAGCTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 72961  
Qy 3297 GGGCTATGCTGTAAATGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 3356  
Db 72960 GGGCTATGCTGTAAATGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 72901  
Qy 3357 GCAATTTGAGCAAGTCAAGGAGCAACAGAGAGCCCATATCTAATAAAACAAACAA 3416  
Db 72900 GCAATTTGAGCAAGTCAAGGAGCAACAGAGAGCCCATATCTAATAAAACAAACAA 72841  
Qy 3417 ACATAATTAACCTGAGTATGATGATGATGATGATGATGATGATGATGATG 3476  
Db 72840 ACATAATTAACCTGAGTATGATGATGATGATGATGATGATGATGATGATG 72781  
Qy 3477 CAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3536  
Db 72780 CAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 72721  
Qy 3537 ACTCAGTCTGAGTGAAGTGAAGAGTGTCTCAAAAATTAATTAATTAATTAAT 3596  
Db 72720 ACTCAGTCTGAGTGAAGTGAAGAGTGTCTCAAAAATTAATTAATTAATTAAT 72661



OY	3397	TTTTAAAAA	CAAAAATTTA	TTAAATTTT	AAAAACCA	CAACATGAGATGTTG	CAAA	3656		
Db	72660	TTTTAAAAA	CAAAAATTTA	TTAAATTTT	AAAAACCA	CAACATGAGATGTTG	CAAA	72601		
OY	3657	TTGATTAAT	TGGGATCT	TAATCC	TCGGAATTA	TTTAAATATT	TGAAAGATCTTC	3716		
Db	72600	TTGATTAAT	TGGGAGCT	TAATAT	CCCTGGAATTA	TTTAAATATT	TGAAAGATCTTC	72541		
OY	3717	CTCAATTC	CTAGAGAC	GTGCAATT	GTAAATAT	TCAGAGCT	TCAGAGGAA	CACTAGAGGCTCGC	3776	
Db	72540	CTCAATTC	CTAGAGAC	GTGCAATT	GTAAATAT	TCAGAGCT	TCAGAGGAA	CACTAGAGGCTCGC	72481	
OY	3777	CATCCAAAG	TGTGTGCT	CCAGAGAC	CAACACAT	CAAGTA	CCCTGGCAAC	GTGTAGAAAT	3836	
Db	72480	CATCCAAAG	TGTGTGCT	CCAGAGAC	CAACACAT	CAAGTA	CCCTGGCAAC	GTGTAGAAAT	72421	
OY	3837	GCAGATCT	TAAAGCTC	ACCCAGAC	CTA	CTGAA	CCAGATCT	GCATTAACAAATTTCT	3896	
Db	72420	GCAGATCT	TAAAGCTC	ACCCAGAC	CTA	CTGAA	CCAGATCT	GCATTAACAAATTTCT	72361	
OY	3897	AGGTGCTC	ACGGGACA	TTTAAAC	CTTGAGAAC	CTCTGCA	CTAGAAATCT	CACTCACAC	3956	
Db	72360	AGGTGCTC	ACGGGACA	TTTAAAC	CTTGAGAAC	CTCTGCA	CTAGAAATCT	CACTCACAC	72301	
OY	3957	TTTCAATTA	TAAATG	CAATCA	CTTGAGGTG	GTGCA	AAGAAA	TGATATTTTAAATTC	4016	
Db	72300	TTTCAATTA	TAAATG	CAATCA	CTTGAGGTG	GTGCA	AAGAAA	TGATATTTTAAATTC	72242	
OY	4017	AGAACTTT	CTAATTTA	AGTCAT	CTATATTT	CTAATG	CAAGGAA	AGAAAGCCAAATCTCTT	4076	
Db	72241	AGAACTTT	CTAATTTA	AGTCAT	CTATATTT	CTAATG	CAAGGAA	AGAAAGCCAAATCTCTT	72182	
OY	4077	AACGTCAATTA	CAAAATCT	ATAATTA	ATTAGTA	TAGCAATCT	CCCTTAAAG	TTTAAATCAAT	4136	
Db	72181	AACGTCAATTA	CAAAATCT	ATAATTA	ATTAGTA	TAGCAATCT	CCCTTAAAG	TTTAAATCAAT	72122	
OY	4137	TTTGTGAC	ACAGAC	GTGTTGA	TTTGAGCT	GGGGCT	CAAGCG	CGGTGTTGAAATTCAC	4196	
Db	72121	TTTGTGAC	ACAGAC	GTGTTGA	TTTGAGCT	GGGGCT	CAAGCG	CGGTGTTGAAATTCAC	72062	
OY	4197	AATTCACAGAT	TGTAAGCT	GTCTG	GGCTAAG	TAAAGAA	AGAAATGCA	AGTTTAAAT	4256	
Db	72061	AATTCACAGAT	TGTAAGCT	GTCTG	GGCTAAG	TAAAGAA	AGAAATGCA	AGTTTAAAT	72002	
OY	4257	AGCTTCTCC	CTTCAT	CCTG	CTGAG	AGCA	CAAA	TTAAATTTT	TATGAAACACATTT	4316
Db	72001	AGCTTCTCC	CTTCAT	CCTG	CTGAG	AGCA	CAAA	TTAAATTTT	TATGAAACACATTT	71942
OY	4317	AGGTAGATTA	CTTA	CAAGGAAA	ATGCAAA	TTTCTG	GAAGGGCT	TATGATTTGCTCA	4376	
Db	71941	AGGTAGATTA	CTTA	CAAGGAAA	ATGCAAA	TTTCTG	GAAGGGCT	TATGATTTGCTCA	71882	
OY	4377	CAACTTGA	CAATCTA	GTAGATG	CACTA	TTTACAG	GTGTGCT	GTGTAAGG	GGGTGTA	4436
Db	71881	CAACTTGA	CAATCTA	GTAGATG	CACTA	TTTACAG	GTGTGCT	GTGTAAGG	GGGTGTA	71822
OY	4437	GGGAGATG	TGAATCA	CAACAT	GTATG	AGACGG	TATGAT	CAACAG	GTGTGTTTTTTTCCC	4496
Db	71821	GGGAGATG	TGAATCA	CAACAT	GTATG	AGACGG	TATGAT	CAACAG	GTGTGTTTTTTTCCC	71762
OY	4497	CCGTGTGA	GTCTAAT	CCCAATG	AGCTT	CTGAA	TCAATTT	CATTCAATTTCCAAATCCA	4556	
Db	71761	CCGTGTGA	GTCTAAT	CCCAATG	AGCTT	CTGAA	TCAATTT	CATTCAATTTCCAAATCCA	71702	
OY	4557	CAAAAC	CAAGAT	TAACTT	TACAGCC	CAATTC	CAAGAA	AGGAATTAATATT	TTTGTGTGAG	4616
Db	71701	CAAAAC	CAAGAT	TAACTT	TACAGCC	CAATTC	CAAGAA	AGGAATTAATATT	TTTGTGTGAG	71642
OY	4617	ACTTTC	CTGATATTA	CACTGA	TTTGGGAA	TATATG	CAAA	TTTTATGTTCTCCTTGGA	4676	
Db	71641	ACTTTC	CTGATATTA	CACTGA	TTTGGGAA	TATATG	CAAA	TTTTATGTTCTCCTTGGA	71582	
OY	4677	GTAGGTCA	AAATGCA	CAAAA	CCAAAA	CAGCAAAA	CTGTAA	CACTAAAGATTAAGCT	4736	

[illegible]



#P 30 11:51:28 2004

UB-10-030-294-2.Std.rge

Page 9

LOCUS 70501 ACATCCAGCGGTGAGTTGGAATGAGACATTAATCTTCTCAAAACCTTAATGAGTGCCTT 70442  
QY 5817 GTGATATTAAGATGCTGACGCTGTGTACAAAGAAACAATTCCTCTCTTGTATAGTCAGCA 5876  
DB 70441 GGTGATTAAGATGCTGACGCTGTGTACAAAGAAACAATTCCTCTCTTGTATAGTCAGCA 70382  
QY 5877 GTGATATTAATTTGAGACTTCTGTGACTTAAGAGTGTGTGAGACATATTTTCTGATG 5936  
DB 70381 GTGATATTAATTTGAGACTTCTGTGACTTAAGAGTGTGTGAGACATATTTTCTGATG 70322  
QY 5937 TCTTTTTCGTTGATATTTGATC 5960  
DB 70321 TCTTTTTCGTTGATATTTGATC 70298

RESULT 3  
LOCUS BD093100 13940 bp DNA linear PAT 27-AUG-2002  
DEFINITION Gene encoding promoter region of tumor suppressor gene p51.  
ACCESSION BD093100  
VERSION BD093100.1 GI:22638688  
KEYWORDS MO 0100818-A/12.  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.  
1 (bases 1 to 13940)  
REFERENCE  
AUTHORS Sakai, T., Kagaya, S., Sato, T., Sukekaga, Y. and Fujii, H.  
TITLE Gene encoding promoter region of tumor suppressor gene p51  
JOURNAL Patent: WO 0100818-A 12 04-JAN-2001;  
TOSHIYUKI SAKAI, NIPPON KAYAKU KK, SHIGEHIDE KAGAYA, TAKAMICHI SATO, YOSHIKAZU  
YOSHIKAZU SUKENAGA, HIDEJI FUJII  
COMMENT  
OS Artificial Sequence  
PN MO 0100818-A/12  
PD 04-JAN-2001  
PF 28-JUN-2000 MO 2000JP004261  
PR 29-JUN-1999 JP 99 183195  
PI TOSHIYUKI SAKAI, SHIGEHIDE KAGAYA, TAKAMICHI SATO, YOSHIKAZU PI  
SUKENAGA, HIDEJI FUJII  
PC C12N15/10, C12N5/10, C12Q1/68, A61K48/00, A61K45/00, A61P43/00, PC  
A61P35/00  
CC Plasmid gene containing p51promoter and neomycin resistance  
CC  
FH Key Location/Qualifiers.  
FEATURES  
source Location/Qualifiers  
1..13940  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

ORIGIN  
Query Match 95.3%; Score 5681.6; DB 6; Length 13940;  
Best Local Similarity 99.3%; Pred. No. 0; Mismatches 39; Indels 0; Gaps 0;  
Matches 5705; Conservative 0; Mismatches 39; Indels 0; Gaps 0;  
QY 2 AGCTGTTGAGGATGCTGGAAGAAAGAAAGCCCAACCAATTCCTTGTGACACTGGGTG 61  
DB 51 ATCTGTTGAGGATGCTGGAAGAAAGAAAGCCCAACCAATTCCTTGTGACACTGGGTG 110  
QY 62 ACTTGGAGGATGCTGAGTGTCTGTTAAGAAATGCGCAACCTCTTCTGCCCAAT 121  
DB 111 ACTTGGAGGATGCTGAGTGTCTGTTAAGAAATGCGCAACCTCTTCTGCCCAAT 170  
QY 122 GGCCTGCTGCTGCTGAGAGCGCTCTTCTGAGGACCTGCTTAAGGACATCTTCTTA 181  
DB 171 GGCCTGCTGCTGCTGAGAGCGCTCTTCTGAGGACCTGCTTAAGGACATCTTCTTA 230  
QY 182 CATTAAGTAATTAATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 241  
DB 231 CATTAAGTAATTAATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 290  
QY 242 TCATAATGAGAGAGCAATGCTAAGTTGACATGCACTTCTTTTTCATTAATTAATTA 301

DB 291 TCATAATGAGAGAGCAATGCTAAGTTGACATGCACTTCTTTTTCATTAATTAATTA 350  
QY 302 GTCAATGAAATACCAATTCAGTTACTTAAGTTCTAGGCCCCAGCTTACTCTAATGATG 361  
DB 351 GTCAATGAAATACCAATTCAGTTACTTAAGTTCTAGGCCCCAGCTTACTCTAATGATG 410  
QY 362 TCAGACTGAGCAATATTAAGTCCAAAGCTGGAAGGATTAAGAGATCTCTCCATGAC 421  
DB 411 TCAGACTGAGCAATATTAAGTCCAAAGCTGGAAGGATTAAGAGATCTCTCCATGAC 470  
QY 422 AGAATCTTGCTCCACTTACTTAATAAATAGATTTGTGTGTTAGCTGAGCTATGAC 481  
DB 471 AGAATCTTGCTCCACTTACTTAATAAATAGATTTGTGTGTTAGCTGAGCTATGAC 530  
QY 482 AGAAGCTGACATCAATTAATAAATACCAAACTGAGCTCTCAATTTAGCAATGAT 541  
DB 531 AGAAGCTGACATCAATTAATAAATACCAAACTGAGCTCTCAATTTAGCAATGAT 590  
QY 542 GGTGAATTTACCCACCACTCTCTCTGTAAGGTTCTTCTGCTCTCTCACTATAA 601  
DB 591 GGTGAATTTACCCACCACTCTCTCTGTAAGGTTCTTCTGCTCTCTCACTATAA 650  
QY 602 ATGCAAGATGACCTGGAAGAGCTAGACCTGAGTTCAATTACCTGACACAAAGAAAT 661  
DB 651 ATGCAAGATGACCTGGAAGAGCTAGACCTGAGTTCAATTACCTGACACAAAGAAAT 710  
QY 662 CAGTTTCTGATCTCAATGTCAGAGCTGCGAGGCTTAAGGACATGCAAGATGAT 721  
DB 711 CAGTTTCTGATCTCAATGTCAGAGCTGCGAGGCTTAAGGACATGCAAGATGAT 770  
QY 722 TCGTTTAAGCTCTGAGTGTGATCTGTGTTTTCACAGTCCCTGTAACCTATTTGCTT 781  
DB 771 TCGTTTAAGCTCTGAGTGTGATCTGTGTTTTCACAGTCCCTGTAACCTATTTGCTT 830  
QY 782 TCGTTTGTGAAGAACCTTATTTCTCTGAAATCTCTGCTCAGTCAATGATGAGG 841  
DB 831 TCGTTTGTGAAGAACCTTATTTCTCTGAAATCTCTGCTCAGTCAATGATGAGG 890  
QY 842 CATCAATGACATGATGAGGCTCTCTGCGCCAAACATGACATCTTCTTTGGGAAT 901  
DB 891 CATCAATGACATGATGAGGCTCTCTGCGCCAAACATGACATCTTCTTTGGGAAT 950  
QY 902 GAACTTAAGCTGAATGCTGAAGTTCAAAAAAGCTGTGAATGATGATGAGCTTACA 961  
DB 951 GAACTTAAGCTGAATGCTGAAGTTCAAAAAAGCTGTGAATGATGATGAGCTTACA 1010  
QY 962 GTGCTTGGCAAGTGAAGTCAATTCCTAATTTCTTAAGTCCCTGAATTTAATTAATC 1021  
DB 1011 GTGCTTGGCAAGTGAAGTCAATTCCTAATTTCTTAAGTCCCTGAATTTAATTAATC 1070  
QY 1022 CTGTTACAGCCCTTGTGAGATGTGTGTTTTTTCACATGCTCTCTAATGCTGTG 1081  
DB 1071 CTGTTACAGCCCTTGTGAGATGTGTGTTTTTTCACATGCTCTCTAATGCTGTG 1130  
QY 1082 AATTTTCAATTTCTTCTTCAATCAATTTCAATTTCTTCTTCTTCTTCTTCTTCTT 1141  
DB 1131 AATTTTCAATTTCTTCTTCAATCAATTTCAATTTCTTCTTCTTCTTCTTCTTCTT 1190  
QY 1142 GCTTAAAGTGAAGCAATCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1201  
DB 1191 GCTTAAAGTGAAGCAATCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1250  
QY 1202 CCTTCCCTTTAAATAATTAAGTATCTAAGGCTCAAAAGAGTGAAGCTCTGAGCT 1261  
DB 1251 CCTTCCCTTTAAATAATTAAGTATCTAAGGCTCAAAAGAGTGAAGCTCTGAGCT 1310  
QY 1262 CTGAGAGTAAGTGAATCAAGAGCTGTAATCAATCCCAAGTTAGCAATGCTTTAACA 1321  
DB 1311 CTGAGAGTAAGTGAATCAAGAGCTGTAATCAATCCCAAGTTAGCAATGCTTTAACA 1370  
QY 1322 ACATTTGCTCTCTCTGAGTGAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1361  
DB 1371 ACATTTGCTCTCTCTGAGTGAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1430

BEST AVAILABLE COPY

QY 1382 AAAGACCTCAGAGGAGCAAGAAAGATTCTGGGATATGTGATGATGAACTCTTAATAA 1441  
DB 1431 AAAGACCTCAGAGGAGCAAGAAAGATTCTGGGATATGTGATGATGAACTCTTAATAA 1490  
QY 1442 GTTTCGTGAGCTCTGGCCCTAATTTGTGATCTTAAGACCAATTCATTTCTTAATAGCTA 1501  
DB 1491 GTTTCGTGAGCTCTGGCCCTAATTTGTGATCTTAAGACCAATTCATTTCTTAATAGCTA 1550  
QY 1502 AACAAACAAACAGAGATCCAGAGTTCAGCAGCTAATAAAGATGAATTAATCTGATACA 1561  
DB 1551 AACAAACAAACAGAGATCCAGAGTTCAGCAGCTAATAAAGATGAATTAATCTGATACA 1610  
QY 1562 GTTGACATCATGATTAATCTGAGAAACATGGGATCAATAGCAAAAAATTCAGACA 1621  
DB 1611 GTTGACATCATGATTAATCTGAGAAACATGGGATCAATAGCAAAAAATTCAGACA 1670  
QY 1622 CAGAAAGAAATGATACATATGCTGATTTATGTGATATTCAGCATTTGATGTCACAC 1681  
DB 1671 CAGAAAGAAATGATACATATGCTGATTTATGTGATATTCAGCATTTGATGTCACAC 1730  
QY 1682 ATATAGCCAGATACCTCAGATGCTAATCAATTTAGTTGATTAATAATTAAGTAAGA 1741  
DB 1731 ATATAGCCAGATACCTCAGATGCTAATCAATTTAGTTGATTAATAATTAAGTAAGA 1790  
QY 1742 TAAAAATTTAGCTCTTCAGATAGGCTTAAAGCAGATGTAATCTAGGCTACACATCAGACG 1801  
DB 1791 TAAAAATTTAGCTCTTCAGATAGGCTTAAAGCAGATGTAATCTAGGCTACACATCAGACG 1850  
QY 1802 GTGCAATATGAAATATTTCTTTATTAACAAAGATCTATTTGCAAAACAAATGTTCTAGA 1861  
DB 1851 GTGCAATATGAAATATTTCTTTATTAACAAAGATCTATTTGCAAAACAAATGTTCTAGA 1910  
QY 1862 AATATATACATATCTATTAATAAACAAGAGATGAGATGTTCTAAGGCGAGGGGT 1921  
DB 1911 AATATATACATATCTATTAATAAACAAGAGATGAGATGTTCTAAGGCGAGGGGT 1970  
QY 1922 GAGGGAGATGATGATGCAAAAGGTGATATGAGAAAGTTTGGGGTAATAGGCTTTTGGA 1981  
DB 1971 GAGGGAGATGATGATGCAAAAGGTGATATGAGAAAGTTTGGGGTAATAGGCTTTTGGA 2030  
QY 1982 TCTTGATTTGCGATGAAAGGCTACTCGGTGCTAATGTGTCACTCTCAGACATGAACTT 2041  
DB 2031 TCTTGATTTGCGATGAAAGGCTACTCGGTGCTAATGTGTCACTCTCAGACATGAACTT 2090  
QY 2042 GGAATGCGCAATTTCAATGTAATGTAATTAATTAACCTCAATAAGTAATCTTAAGAGTGA 2101  
DB 2091 GGAATGCGCAATTTCAATGTAATGTAATTAATTAACCTCAATAAGTAATCTTAAGAGTGA 2150  
QY 2102 GTCTTTTGTGCAAAATTAATTTTAATCACTGCAATTAATTAATGATGATTTTGTGCA 2161  
DB 2151 GTCTTTTGTGCAAAATTAATTTTAATCACTGCAATTAATTAATGATGATTTTGTGCA 2210  
QY 2162 ATACATTAACATCTATTAATCAATTAAGTGCATAATTTTGTGCTGCTGCAAAACCA 2221  
DB 2211 ATACATTAACATCTATTAATCAATTAAGTGCATAATTTTGTGCTGCTGCAAAACCA 2270  
QY 2222 GAGATGAATTAAGATGCAAGGCTGCTGCAAGGCTTTCAATTCAGATCAAGAGATGA 2281  
DB 2271 GAGATGAATTAAGATGCAAGGCTGCTGCAAGGCTTTCAATTCAGATCAAGAGATGA 2330  
QY 2282 GTCTATTTCAAAATTAAGCTAGCTCCAGAGAAAGTTATAGTGAATTCACAAAAAAG 2341  
DB 2331 GTCTATTTCAAAATTAAGCTAGCTCCAGAGAAAGTTATAGTGAATTCACAAAAAAG 2390  
QY 2342 TGCAATATTAATTAATGAGACAGTGAAGATGAGGAAAGTTTCTTTTAATGTTGAAAAAAG 2401  
DB 2391 TGCAATATTAATTAATGAGACAGTGAAGATGAGGAAAGTTTCTTTTAATGTTGAAAAAAG 2450  
QY 2402 AGGAGAAATTTTGTGCTTTGAAGATGAAGATGAAGATGAATATGCGCAGATGAGATT 2461  
DB 2451 AGGAGAAATTTTGTGCTTTGAAGATGAAGATGAAGATGAATATGCGCAGATGAGATT 2510

QY 2462 AAAACATTCCTGTGTGAGGGGAGAAATATGATCCAGGCAAGAGCAACAGAAAAATAT 2521  
DB 2511 AAAACATTCCTGTGTGAGGGGAGAAATATGATCCAGGCAAGAGCAACAGAAAAATAT 2570  
QY 2522 GCACCTTAGAGAAAGTATGAAAGGGGAGCAGTTGTAATTAATTTCAATGAAATGTAAG 2581  
DB 2571 GCACCTTAGAGAAAGTATGAAAGGGGAGCAGTTGTAATTAATTTCAATGAAATGTAAG 2630  
QY 2582 TGAGAAAGATTTGATCAATGACACCTGAGTTGGCAGATGCTAGTTGCTGCTAG 2641  
DB 2631 TGAGAAAGATTTGATCAATGACACCTGAGTTGGCAGATGCTAGTTGCTGCTAG 2690  
QY 2642 GAGTCAGAGAAACAAAGTGTCCCTTTCTCTACGTTATGCTCAGTGTCCAGTCCAAA 2701  
DB 2691 GAGTCAGAGAAACAAAGTGTCCCTTTCTCTACGTTATGCTCAGTGTCCAGTCCAAA 2750  
QY 2702 AACCTTTCTCTCTTAAGTACTTTCTTCCCTCCATACAAATCTTAATCTTTCA 2761  
DB 2751 AACCTTTCTCTCTTAAGTACTTTCTTCCCTCCATACAAATCTTAATCTTTCA 2810  
QY 2762 ACATCATTTAAACAGGAGGTCATGTGCAAGAAAGCAATTCCTTTCTTACACTTATG 2821  
DB 2811 ACATCATTTAAACAGGAGGTCATGTGCAAGAAAGCAATTCCTTTCTTACACTTATG 2870  
QY 2822 TACGTTATTAATATACAAATTTCTGCTAAAGACTCTAAAGCTTGGAAAAAGTTTCAC 2881  
DB 2871 TACGTTATTAATATACAAATTTCTGCTAAAGACTCTAAAGCTTGGAAAAAGTTTCAC 2930  
QY 2882 TTGCACATCAAGATATTAATTCATGATGTTGTATGTAACCTTACCTCCCTAAGAAATA 2941  
DB 2931 TTGCACATCAAGATATTAATTCATGATGTTGTATGTAACCTTACCTCCCTAAGAAATA 2990  
QY 2942 AGATGAACATTAATTAATTAAGAAATTAATGTAATTAATTAATGATGCACTTAAT 3001  
DB 2991 AGATGAACATTAATTAATTAAGAAATTAATGTAATTAATTAATGATGCACTTAAT 3050  
QY 3002 TTCACTTGATGCTATGATGTTGCAATGCTACTGCTGTGTTGTAATCTAAGAGTTTG 3061  
DB 3051 TTCACTTGATGCTATGATGTTGCAATGCTACTGCTGTGTTGTAATCTAAGAGTTTG 3110  
QY 3062 CTCTTTTCTCGGTGCACTCTGCAATTAATTTGCAATATGCAATGAGCAATCTTT 3121  
DB 3111 CTCTTTTCTCGGTGCACTCTGCAATTAATTTGCAATATGCAATGAGCAATCTTT 3170  
QY 3122 TCAATATTAATTAATTAATTTGTCATTTAATTTCTGTTCTCTTACCTTATGTA 3181  
DB 3171 TCAATATTAATTAATTAATTTGTCATTTAATTTCTGTTCTCTTACCTTATGTA 3230  
QY 3182 CTTTGAATTTTAAATTAACAACATATGAAATCAATGACATAGCTTTAAATGATTAAT 3241  
DB 3231 CTTTGAATTTTAAATTAACAACATATGAAATCAATGACATAGCTTTAAATGATTAAT 3290  
QY 3242 AAATAGCTTGAATTAATTAATCTTTAAATTTTAAATAATTAATGATGATGTTGCTC 3301  
DB 3291 AAATAGCTTGAATTAATTAATCTTTAAATTTTAAATAATTAATGATGATGTTGCTC 3350  
QY 3302 ATGCTGTAAATCCCAACATCTTTGGAAGCCGGGTGCGGAGGATAGTGTGATGCTCAG 3361  
DB 3351 ATGCTGTAAATCCCAACATCTTTGGAAGCCGGGTGCGGAGGATAGTGTGATGCTCAG 3410  
QY 3362 TTGAGACCACTGAGGCAACAGAGAACCCATATTTTAAATAATTAATGATGATGTTGCTC 3421  
DB 3411 TTGAGACCACTGAGGCAACAGAGAACCCATATTTTAAATAATTAATGATGATGTTGCTC 3470  
QY 3422 ATTACCTGGATGATGTTGCTCACTGATGCTCAAGCTACAGAGAAAGCTGAGGAGAA 3481  
DB 3471 ATTACCTGGATGATGTTGCTCACTGATGCTCAAGCTACAGAGAAAGCTGAGGAGAA 3530  
QY 3482 GGAATCACTTGAAGCCAGAGAGTTGAGCTGAGAGATCAATGAAGCCGCTGATCACTCA 3541  
DB 3531 GGAATCACTTGAAGCCAGAGAGTTGAGCTGAGAGATCAATGAAGCCGCTGATCACTCA 3590  
QY 3542 CTCGCTGACAGTGAAGAGAGCTGTCTCAAAAATTAATTAATTAATTAATTAATCTTTA 3601

Db	4671	CCGAAATTAACACGATGATTTGGGAATATATGACAAATTTATATGATTTCCCTTCGAAAGTAGC	4730
Oy	4682	TCGAGTCGAAAGCAAAACCAAAAACAGCAAAAACCTGTAACATTAAGATAGATGAGC	4741
Db	4731	TCGAGTCGAAAGCAAAACCAAAAACAGCAAAAACCTGTAACATTAAGATAGATGAGC	4790
Oy	4742	CGACTGAGAGATTAATAAATAAATAAGATATTTTATTAACAGGCATTTGAATATTTG	4801
Db	4791	CGACTGAGAGATTAATAAATAAATAAGATATTTTATTAACAGGCATTTGAATATTTG	4850
Oy	4802	TGCACTCGAATATTTCAACAATTAATATATTTTCAATTTATATCTTTAAGAAAT	4861
Db	4851	TGCACTCGAATATTTCAACAATTAATATATTTTCAATTTATATCTTTAAGAAAT	4910
Oy	4862	TACTATATATATATGTAACATCTGCAATGCTTGAAGTACGATATTTAACTCAATPAA	4921
Db	4911	TACTATATATATATGTAACATCTGCAATGCTTGAAGTACGATATTTAACTCAATPAA	4970
Oy	4922	GGTATATTTCTTTATTTGCGGTCAAGCAAGCTTCTAAAGGAGATGTGAAAGGATATCTC	4981
Db	4971	GGTATATTTCTTTATTTGCGGTCAAGCAAGCTTCTAAAGGAGATGTGAAAGGATATCTC	5030
Oy	4982	TTTTCTTTAGCTGAAGAAAGAGAGATTTCTTAAGTAAATATATCAAGAAATTTCCCTG	5041
Db	5031	TTTTCTTTAGCTGAAGAAAGAGAGATTTCTTAAGTAAATATATCAAGAAATTTCCCTG	5090
Oy	5042	TCTTTGCTATTTGAGATTTGAGCACAACAGCGATTTGCTGAAAGGAACTGAAGGC	5101
Db	5091	TCTTTGCTATTTGAGATTTGAGCACAACAGCGATTTGCTGAAAGGAACTGAAGGC	5150
Oy	5102	GGGAGAGGAGGAAATATGATGAATAAAACAAAACAAAACCTTCCCTAAGCAGCTCTA	5161
Db	5151	GGGAGAGGAGGAAATATGATGAATAAAACAAAACAAAACCTTCCCTAAGCAGCTCTA	5210
Oy	5162	CAAAACATTTTGGCCCCGAAATATGTCACGAAATCTCAAAATCAACACGATATCCAGAT	5221
Db	5211	CAAAACATTTTGGCCCCGAAATATGTCACGAAATCTCAAAATCAACACGATATCCAGAT	5270
Oy	5222	ACGAGGAGTGTATGTATGTAGCTGAGCAAGGAGTGAGCACTCATCAGCTCAGTTCAATTCAA	5281
Db	5271	ACGAGGAGTGTATGTATGTAGCTGAGCAAGGAGTGAGCACTCATCAGCTCAGTTCAATTCAA	5330
Oy	5282	AGGTCCAGGCTGCTGAATTTAACTCTGATATCCATTCATCCAGCTATCCATCAAGCAG	5341
Db	5331	AGGTCCAGGCTGCTGAATTTAACTCTGATATCCATTCATCCAGCTATCCATCAAGCAG	5390
Oy	5342	AGATCAGAGTTGACAGATGCTCTCAAGTCCAAATTTGCCAAAGAAAGATGAGCTATAT	5401
Db	5391	AGATCAGAGTTGACAGATGCTCTCAAGTCCAAATTTGCCAAAGAAAGATGAGCTATAT	5450
Oy	5402	ACGTCAAGCACTCTGAAGCCGCTGAGAGAGGGGAAAGAACAGATGAGAGGATGCCAG	5461
Db	5451	ACGTCAAGCACTCTGAAGCCGCTGAGAGAGGGGAAAGAACAGATGAGAGGATGCCAG	5510
Oy	5462	CTGGTAGAGATCGAAGTGTATGACAGTTTATAGTCAATTGATGATATCTCATTTGGCTAAAT	5521
Db	5511	CTGGTAGAGATCGAAGTGTATGACAGTTTATAGTCAATTGATGATATCTCATTTGGCTAAAT	5570
Oy	5522	CAAGAAACGCTCCGCTCTTTGCAATATGATGACAGAGAACAGTCCCTAACTTTAT	5581
Db	5571	CAAGAAACGCTCCGCTCTTTGCAATATGATGACAGAGAACAGTCCCTAACTTTAT	5630
Oy	5582	GTCTGATAGCAATTTGAACCTATTTGCTTTTACGCTCCGACTTTATATCTATATATACAA	5641
Db	5631	GTCTGATAGCAATTTGAACCTATTTGCTTTTACGCTCCGACTTTATATCTATATATACAA	5680
Oy	5642	GGATATTTGATATATTTATATATATTTGTTCCGCTCGTCAATATCAAGACAGTTTAG	5701
Db	5691	GGATATTTGATATATTTATATATATTTGTTCCGCTCGTCAATATCAAGACAGTTTAG	5750
Oy	5702	GAATATATTTTGAACCTCACAGTGTGCAACCCCAAGCAAGTACAG	5745
Db	5751	TAAGACGCGCTCGAGATCTTAAGTAAACCTTGCGATCTTCGATACAG	5794

QY	481	CAGAAAGCTCAATCAATTAATAATGCAAAATCTGAGTCTTCAATTTGACCAATAGT	540
Db	481	CAGAAAGTCTCAATCAATTAATAATAATGCAAAATCTGAGTCTTCAATTTGACCAATAGT	540
QY	541	TGTGAAATTAATCTCCACCACTCTCTCTTGAAAGTTCTTCTGCTCTCTCACTATA	600
Db	541	TGTGAAATTAATCTCCACCACTCTCTCTTGAAAGTTCTTCTGCTCTCTCACTATA	600
QY	601	AATGCAAGATGACCTGAAAGGCTGAGACTGAGGTTCACTAACCTGACACAAAGAAAT	660
Db	601	AATGCAAGATGACCTGAAAGGCTGAGACTGAGGTTCACTAACCTGACACAAAGAAAT	660
QY	661	TGAGTTTCTCATCTCCATAGTCCACAGGCTCCGAGCTCTAAGGAACTGCAAGATCAT	720
Db	661	TGAGTTTCTCATCTCCATAGTCCACAGGCTCCGAGCTCTAAGGAACTGCAAGATCAT	720
QY	721	CTGCTTAAAGCTCTTGTGTGTGACATCTGTGTGTTTCCACTGACCTGTACCTATGCTCT	780
Db	721	CTGCTTAAAGCTCTTGTGTGTGACATCTGTGTGTTTCCACTGACCTGTACCTATGCTCT	780
QY	781	TTCCCTGTTAACAAGACCTTATATTTCTCTGAAAATCTCTGCTCAGTCAATGAGAGG	840
Db	781	TTCCCTGTTAACAAGACCTTATATTTCTCTGAAAATCTCTGCTCAGTCAATGAGAGG	840
QY	841	CCATCACTCAACATGATGAGGCTCTCTGTGGCAAACTGACATCTTTTGTGGAAAT	900
Db	841	CCATCACTCAACATGATGAGGCTCTCTGTGGCAAACTGACATCTTTTGTGGAAAT	900
QY	901	TGAATCTTAAAGCTGAATGACTGAAGTCAAAAAAGCTGTGTAATCTGACTTAAGCTTAC	960
Db	901	TGAATCTTAAAGCTGAATGACTGAAGTCAAAAAAGCTGTGTAATCTGACTTAAGCTTAC	960
QY	961	AGTGGCTTTGCAAGATGACTGTCCATCTCTATCTTAACTGCTCAATTTTAAATTAAT	1020
Db	961	AGTGGCTTTGCAAGATGACTGTCCATCTCTATCTTAACTGCTCAATTTTAAATTAAT	1020
QY	1021	CTGTGTTACAGGCTTTCTGAGAATGTGTGTTTTTTCCAACTGTCTCTAATAGTCTGT	1080
Db	1021	CTGTGTTACAGGCTTTCTGAGAATGTGTGTTTTTTCCAACTGTCTCTAATAGTCTGT	1080
QY	1081	GAAATTTCAATTTCTTTTCAATCTTTCATGTTTGTGTGTTGTGTGTGTGTGT	1140
Db	1081	GAAATTTCAATTTCTTTTCAATCTTTCATGTTTGTGTGTTGTGTGTGTGTGT	1140
QY	1141	GGGCTTTAGGTAGGCAATCAATGTTCTGTGTGTTATACCCAAAGAAATCTGATATATCA	1200
Db	1141	GGGCTTTAGGTAGGCAATCAATGTTCTGTGTGTTATACCCAAAGAAATCTGATATATCA	1200
QY	1201	TTCTTCCCTTTTAAAAATTAAGTATCTAAAGCTCAAGAGATGAGGCTACCTGCTGAG	1260
Db	1201	TTCTTCCCTTTTAAAAATTAAGTATCTAAAGCTCAAGAGATGAGGCTACCTGCTGAG	1260
QY	1261	TCTGGAGTAAAGTAAAGTACCAAGACTGTACTAACCCCAAGTTAAGCTCACTGCTTAAAC	1320
Db	1261	TCTGGAGTAAAGTAAAGTACCAAGACTGTACTAACCCCAAGTTAAGCTCACTGCTTAAAC	1320
QY	1321	AACATTTGCTCTCTCTCTCAAGTTAATAGCAAGTTTGGAGAAAGAAAGCTACATTTTGG	1380
Db	1321	AACATTTGCTCTCTCTCTCAAGTTAATAGCAAGTTTGGAGAAAGAAAGCTACATTTTGG	1380
QY	1381	CAAAAGCTCAAGAGACCAAGAACTTGTGGAAATGTGTATCTTGAATCTTTAAAT	1440
Db	1381	CAAAAGCTCAAGAGACCAAGAACTTGTGGAAATGTGTATCTTGAATCTTTAAAT	1440
QY	1441	AGTTTGTGAACTTTGTGCTCAATTTGTGTATCTAAGACCAAGTTTCAATTTCTTAAATAGT	1500
Db	1441	AGTTTGTGAACTTTGTGCTCAATTTGTGTATCTAAGACCAAGTTTCAATTTCTTAAATAGT	1500
QY	1501	AAAACAAACAAACAAAGATCCACAGCTTGACAGCTATAATTAAGAGTGAATTAATGATAC	1560
Db	1501	AAAACAAACAAACAAAGATCCACAGCTTGACAGCTATAATTAAGAGTGAATTAATGATAC	1560

p 30 11:51:28 2004

us-10-030-294-2.std.rge

Page 13

1561 AGTTGACACATGATATATATCTGAGAACCTGCGATCATATGACCAAAAAATCCGAC 1620  
1561 AGTTGACACATGATATATATCTGAGAACCTGCGATCATATGACCAAAAAATCCGAC 1620  
1561 ACAGAGATACGATACATATGCGCTGATTTATGATATCTAGCATGATTTGTCGA 1680  
1561 ACAGAGATACGATACATATGCGCTGATTTATGATATCTAGCATGATTTGTCGA 1680  
1681 CATAGTACCGATGCTGACATGCTGATTTAAATTTAAATTTAAATTTAAATTTAA 1740  
1681 CATAGTACCGATGCTGACATGCTGATTTAAATTTAAATTTAAATTTAAATTTAA 1740  
1741 ATAAATTTAGCTCTTCACTGATGCTGATTTAAATTTAAATTTAAATTTAAATTT 1800  
1741 ATAAATTTAGCTCTTCACTGATGCTGATTTAAATTTAAATTTAAATTTAAATTT 1800  
1801 GGTGCAATATATGATATTTCTTTATTAACAGAAAGTTCTATTTGCAAAATTTCTAG 1860  
1801 GGTGCAATATATGATATTTCTTTATTAACAGAAAGTTCTATTTGCAAAATTTCTAG 1860  
1861 AAAATATACATATATCTATTAACAGAAAGTTCTATTTGCAAAATTTCTAG 1920  
1861 AAAATATACATATATCTATTAACAGAAAGTTCTATTTGCAAAATTTCTAG 1920  
1921 TGAGGAGATGATGATGCAAGGATGATGATGATGATGATGATGATGATGATGATG 1980  
1921 TGAGGAGATGATGATGCAAGGATGATGATGATGATGATGATGATGATGATGATG 1980  
1981 ATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
1981 ATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
1981 ATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
2041 TGCAATTTGCGAATTTCTATTTATGATGATGATGATGATGATGATGATGATG 2100  
2041 TGCAATTTGCGAATTTCTATTTATGATGATGATGATGATGATGATGATGATG 2100  
2101 AGGTGTTGCGAATTTCTATTTATGATGATGATGATGATGATGATGATGATGATG 2160  
2101 AGGTGTTGCGAATTTCTATTTATGATGATGATGATGATGATGATGATGATGATG 2160  
2101 AGGTGTTGCGAATTTCTATTTATGATGATGATGATGATGATGATGATGATGATG 2160  
2161 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220  
2161 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220  
2161 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220  
2221 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280  
2221 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280  
2281 AGCTATTTGCGAATTTCTATTTATGATGATGATGATGATGATGATGATGATG 2340  
2281 AGCTATTTGCGAATTTCTATTTATGATGATGATGATGATGATGATGATGATG 2340  
2341 GTGCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400  
2341 GTGCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400  
2401 GAGGAGATTTTGTGCTTTGATGATGATGATGATGATGATGATGATGATGATG 2460  
2401 GAGGAGATTTTGTGCTTTGATGATGATGATGATGATGATGATGATGATGATG 2460  
2461 GAGGAGATTTTGTGCTTTGATGATGATGATGATGATGATGATGATGATGATG 2520  
2461 GAGGAGATTTTGTGCTTTGATGATGATGATGATGATGATGATGATGATGATG 2520  
2521 TGCAATTTGCGAATTTCTATTTATGATGATGATGATGATGATGATGATGATG 2580  
2521 TGCAATTTGCGAATTTCTATTTATGATGATGATGATGATGATGATGATGATG 2580  
2581 GTGCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640  
2581 GTGCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640  
2641 GAGGAGATTTTGTGCTTTGATGATGATGATGATGATGATGATGATGATGATG 2700

2641 GAGGAGATTTTGTGCTTTGATGATGATGATGATGATGATGATGATGATGATG 2700  
2701 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760  
2701 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760  
2761 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820  
2761 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820  
2821 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880  
2821 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880  
2881 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2940  
2881 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2940  
2941 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000  
2941 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000  
3001 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3060  
3001 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3060  
3061 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120  
3061 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120  
3121 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180  
3121 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180  
3181 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3240  
3181 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3240  
3241 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300  
3241 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300  
3301 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360  
3301 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360  
3361 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420  
3361 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420  
3421 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3480  
3421 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3480  
3481 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3540  
3481 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3540  
3541 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3600  
3541 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3600  
3601 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3660  
3601 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3660  
3661 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3720  
3661 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3720  
3721 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3780  
3721 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3780

Db 3721 TTTCCTAGACGTGGAATGTAATATATCAAGCTAGAGAGAACATAGGGCTGGCACT 3780  
Qy 3781 CCAAGTGTGTCCAAAGACACAGACATCAAGTAACTGGGAAAGGTGTAAGAAATGCG 3840  
Db 3781 CCAAGTGTGTCCAAAGACACAGACATCAAGTAACTGGGAAAGGTGTAAGAAATGCG 3840  
Qy 3841 AGCTTAGGCTGACCCCAAGACCTACTGAACCAATCTGCAATTACAAATTTCTAGGT 3900  
Db 3841 AGCTTAGGCTGACCCCAAGACCTACTGAACCAATCTGCAATTACAAATTTCTAGGT 3900  
Qy 3901 GCTTCACGGGCAATTTAAACCTTGAGAGCTCTGCAATAGAAATCTTCACTCCACCTTC 3960  
Db 3901 GCTTCACGGGCAATTTAAACCTTGAGAGCTCTGCAATAGAAATCTTCACTCCACCTTC 3960  
Qy 3961 ATTATTAATGGAATCACTTGGGCTGTGTCAAGAGAAATTCATATTTTAAATTCGAA 4020  
Db 3961 ATTATTAATGGAATCACTTGGGCTGTGTCAAGAGAAATTCATATTTTAAATTCGAA 4020  
Qy 4021 CCTTCATTTAGTCACTATATTTTGTATATGCAAGGAGAAAGCCAAATCTTTTACT 4080  
Db 4021 CCTTCATTTAGTCACTATATTTTGTATATGCAAGGAGAAAGCCAAATCTTTTACT 4080  
Qy 4081 GCAATTAACAAATCTATATTAATTAATTAATGAGCAATCTTCCCTTAGTTTACATTTG 4140  
Db 4081 GCAATTAACAAATCTATATTAATTAATTAATGAGCAATCTTCCCTTAGTTTACATTTG 4140  
Qy 4141 TGGAGCAAGCTGTTGATTTGGCTGGGCTCAGGCGGCTGTGTGAAATTCACAT 4200  
Db 4141 TGGAGCAAGCTGTTGATTTGGCTGGGCTCAGGCGGCTGTGTGAAATTCACAT 4200  
Qy 4201 CACGATGTTAGCCGCTCTGGGCTTAAGTAAAGAGAAATGCAATTTTAAATAGCT 4260  
Db 4201 CACGATGTTAGCCGCTCTGGGCTTAAGTAAAGAGAAATGCAATTTTAAATAGCT 4260  
Qy 4261 TCTCCCTTCATCTGGCTGAGAGCAATTAATTTTATGAAGCAATTTTGAAT 4320  
Db 4261 TCTCCCTTCATCTGGCTGAGAGCAATTAATTTTATGAAGCAATTTTGAAT 4320  
Qy 4321 TAGATTTACTTAAGAGAAATGTCATATTTCTGAAAGGCTTACATTTCTGCAAC 4380  
Db 4321 TAGATTTACTTAAGAGAAATGTCATATTTCTGAAAGGCTTACATTTCTGCAAC 4380  
Qy 4381 TTGACATCTACTGATGTCACCTATTATACAGTGTCTGTGACTAGGGGCTGAAGGGA 4440  
Db 4381 TTGACATCTACTGATGTCACCTATTATACAGTGTCTGTGACTAGGGGCTGAAGGGA 4440  
Qy 4441 AGATGTAATCTACCAATGTTAGACCGTATGACAGAGGTGTTTTTCCCGG 4500  
Db 4441 AGATGTAATCTACCAATGTTAGACCGTATGACAGAGGTGTTTTTCCCGG 4500  
Qy 4501 TTGAGCTATCTCACTGATGTCATTTCTGAAATCAATTTCAATTTCCAAATCCAAA 4560  
Db 4501 TTGAGCTATCTCACTGATGTCATTTCTGAAATCAATTTCAATTTCCAAATCCAAA 4560  
Qy 4561 ACCGAGTAAGTTTACAGCCCATTTCCAGAAAGAAATTAATTTTGTGTGAGCT 4620  
Db 4561 ACCGAGTAAGTTTACAGCCCATTTCCAGAAAGAAATTAATTTTGTGTGAGCT 4620  
Qy 4621 TCCGATATATACGATTTGGCAATTAAGAAATTTAGGTTTCTTTCGAGTAG 4680  
Db 4621 TCCGATATATACGATTTGGCAATTAAGAAATTTAGGTTTCTTTCGAGTAG 4680  
Qy 4681 GTCAAGCTCAAGCAAAACCAAAACCAAGCAAACTGTAGACATTAAGAAATGAGTAG 4740  
Db 4681 GTCAAGCTCAAGCAAAACCAAAACCAAGCAAACTGTAGACATTAAGAAATGAGTAG 4740  
Qy 4741 CCGAGTAGAGATTAATTAAGTAAATTTTATTAACAGCAATTTGAATTAATTT 4800  
Db 4741 CCGAGTAGAGATTAATTAAGTAAATTTTATTAACAGCAATTTGAATTAATTT 4800  
Qy 4801 GTGCACTTGAGAAATTTCTACAAATTAATTAATTTTCAATTTTAAATCTTTAAGAAA 4860  
Db 4801 GTGCACTTGAGAAATTTCTACAAATTAATTAATTTTCAATTTTAAATCTTTAAGAAA 4860

Qy 4861 TTACTATATTAATGTAAGTACATGTCATGTTTGAAGTATTAATTAATTAATTA 4920  
Db 4861 TTACTATATTAATGTAAGTACATGTCATGTTTGAAGTATTAATTAATTAATTA 4920  
Qy 4921 AGGTATATTTTCTTTTATTTCCGCTGACGCAAGGCTTTAAGGGAGTGTAAAGGATATCT 4980  
Db 4921 AGGTATATTTTCTTTTATTTCCGCTGACGCAAGGCTTTAAGGGAGTGTAAAGGATATCT 4980  
Qy 4981 CTTTCTCTACGTAAGGAGAAAGTCAAGTTTAAATTAATTAATTAATTAATTTCCCT 5040  
Db 4981 CTTTCTCTACGTAAGGAGAAAGTCAAGTTTAAATTAATTAATTAATTAATTTCCCT 5040  
Qy 5041 GTCTTGTATTTGATGATGTCACCAACAGGCGGTTGCTGAAGGAGAAATGGAAGG 5100  
Db 5041 GTCTTGTATTTGATGATGTCACCAACAGGCGGTTGCTGAAGGAGAAATGGAAGG 5100  
Qy 5101 CCGGAGAGGAGGAAATGAGTGAAGAAACAAACAAACAAATCTTCTTACGAGCTCT 5160  
Db 5101 CCGGAGAGGAGGAAATGAGTGAAGAAACAAACAAACAAATCTTCTTACGAGCTCT 5160  
Qy 5161 ACATAACATTTTACGCCCAAGAAATGTCACAGAAATCTCAATCAACAGATCCAGA 5220  
Db 5161 ACATAACATTTTACGCCCAAGAAATGTCACAGAAATCTCAATCAACAGATCCAGA 5220  
Qy 5221 TACAGAAAGTGTATGATGCTGAGCAGGCTGAGCACTCATGAGCTCAGTTCAAGTAC 5280  
Db 5221 TACAGAAAGTGTATGATGCTGAGCAGGCTGAGCACTCATGAGCTCAGTTCAAGTAC 5280  
Qy 5281 AAAGTCCAGGCTGCTGAATTAATCTGATGCTTCAATGCTCAGCATTCATTCAGACA 5340  
Db 5281 AAAGTCCAGGCTGCTGAATTAATCTGATGCTTCAATGCTCAGCATTCATTCAGACA 5340  
Qy 5341 GAGATCAGAAAGTTCAAGATGCTTCCAGCTCCAAATTCACAAACAAAGGTGCTACTA 5400  
Db 5341 GAGATCAGAAAGTTCAAGATGCTTCCAGCTCCAAATTCACAAACAAAGGTGCTACTA 5400  
Qy 5401 TACGTCAAGACTCTGAAGCGGTGAGAGAGGAGAAACAAAGTGAAGAGATGCCA 5460  
Db 5401 TACGTCAAGAGCTCTGAGCGGTGAGAGAGGAGAAACAAAGTGAAGAGATGCCA 5460  
Qy 5461 GCTGTGAAGATGAGGTGTTATTAAGATTTTATGATTAATGATTAATGCTCAATGCTTAA 5520  
Db 5461 GCTGTGAAGATGAGGTGTTATTAAGATTTTATGATTAATGATTAATGCTCAATGCTTAA 5520  
Qy 5521 TCAAGAAACGCTCGCTCTTTCGAATTAATTAAGAGAGAGAGAGAGAGAGAGAGAG 5580  
Db 5521 TCAAGAAACGCTCGCTCTTTCGAATTAATTAAGAGAGAGAGAGAGAGAGAGAGAG 5580  
Qy 5581 TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5640  
Db 5581 TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5640  
Qy 5641 AGGTATTTGTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 5700  
Db 5641 AGGTATTTGTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 5700

RESULT 5  
BD093091/c 5676 bp DNA linear PAR 27-AUG-2002  
LOCUS BD093091  
DEFINITION Gene encoding promoter region of tumor suppressor gene p51.  
ACCESSION BD093091  
VERSION BD093091.1 GI:22638679  
KEYWORDS MO 0100818-A/3.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 (bases 1 to 5676)  
Sakai, T., Kagaya, S., Sato, T., Sukenaga, Y. and Fujii, H.  
Gene encoding promoter region of tumor suppressor gene p51  
Patent: WO 0100818-A 3 04-JAN-2001;

BEST AVAILABLE COPY



BEST AVAILABLE COPY

TOSHIYUKI SAKAI, NIPPON KAYAKU KK, SHIGEHIDE KAGAWA, TAKAMICHI SATO,  
YOSHIMAZU SUKENAGA, HIDEJI FUJII  
OS Homo sapiens (human)  
PN MO 0100818-N/3  
PD 04-JAN-2001  
PE 28-JUN-2000 MO 2000JP004261  
PR 29-JUN-1999 JP 99P 183195  
PI TOSHIYUKI SAKAI, SHIGEHIDE KAGAWA, TAKAMICHI SATO, YOSHIMAZU  
SUKENAGA, HIDEJI FUJII

PC C12N5/10, C12N5/10, C12Q1/68, A61K48/00, A61K45/00, A61P43/00, PC  
A61P35/00  
CC

FEATURES  
source 1. 5676  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 95.2%; Score 5676; DB 6; Length 5676;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 5676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 CAGCTGTGAGGAGTGTCTGGAAGAGCCCAACCAATGCTCTGAGACATGGGTGT 60
Db 5676 CAGCTGTGAGGAGTGTCTGGAAGAGCCCAACCAATGCTCTGAGACATGGGTGT 5617
Oy 61 GACTTTGAGGGTATCAGGTTGTCTGTAAAGAACTGCAACCTTCCTGCCCCCAAT 120
Db 5616 GACTTTGAGGGTATCAGGTTGTCTGTAAAGAACTGCAACCTTCCTGCCCCCAAT 5557
Oy 121 TGGCTCTGCTCCCTTGCAATGCCCTCTTCTTGGGACATCTCCCTTAAGGCACTTTG 180
Db 5556 TGGCTCTGCTCCCTTGCAATGCCCTCTTCTTGGGACATCTCCCTTAAGGCACTTTG 5497
Oy 181 ACATTAACCTTAACATAAATGTTTATTTGATGAATTTCAATGACCTGAAGAGATGAG 240
Db 5496 ACATTAACCTTAACATAAATGTTTATTTGATGAATTTCAATGACCTGAAGAGATGAG 5437
Oy 241 GTCAATCAGAGAGACATAGCTAGCTTACCTTGAATTCATTTGAAATTA 300
Db 5436 GTCAATCAGAGAGACATAGCTAGCTTACCTTGAATTCATTTGAAATTA 5377
Oy 301 AGTCATTCGAATACATTCAGTTTACTTAAGTTCTAGGCCCAAGCTTTACTCTTAATCAT 360
Db 5376 AGTCATTCGAATACATTCAGTTTACTTAAGTTCTAGGCCCAAGCTTTACTCTTAATCAT 5317
Oy 361 GTCACTGTAGGAAATATTAGTCCAAAGTTGAGAGTTAGAGAGATCTCTCATGA 420
Db 5316 GTCACTGTAGGAAATATTAGTCCAAAGTTGAGAGTTAGAGAGATCTCTCATGA 5257
Oy 421 CAGAACTTGGCTTCACATTAATAATAGATTGTGTGTGAGCTGACGTATGTA 480
Db 5256 CAGAACTTGGCTTCACATTAATAATAGATTGTGTGTGAGCTGACGTATGTA 5197
Oy 481 CAGAAAGTGTCAATCAATTAATAATCAACCAACTCACTCTCAATTTGACCAATAGT 543
Db 5196 CAGAAAGTGTCAATCAATTAATAATCAACCAACTCACTCTCAATTTGACCAATAGT 5137
Oy 543 TGTGAATTTACTCCACACACTCTCTCTGGAAGTTGCTTCTGCTGCTCCCATATA 600
Db 5136 TGTGAATTTACTCCACACACTCTCTCTGGAAGTTGCTTCTGCTGCTCCCATATA 5077
Oy 601 AATGACAGATGACCTGGAAGGCTAGAGCTGAGGTTCAATTACCTTGACCAAGGAAT 660
Db 5076 AATGACAGATGACCTGGAAGGCTAGAGCTGAGGTTCAATTACCTTGACCAAGGAAT 5017
Oy 661 TCAATTTCTGATCTCATAGTCAAGGCTGCCAGAGCTTACAGGAACATGCAAGATCAT 720
Db 5016 TCAATTTCTGATCTCATAGTCAAGGCTGCCAGAGCTTACAGGAACATGCAAGATCAT 4957

```

```

Oy 721 CTGCTTAAGCCCTCTGTGTGAGCATGTGTTTTCACAGCTGCTGATCTCTCT 780
Db 4956 CTGCTTAAGCCCTCTGTGTGAGCATGTGTTTTCACAGCTGCTGATCTCTCT 4897
Oy 781 TTCTTTGTTAAAGAACTTATTTCTCTGAAAACCTCTGCTCAGTCAATGAGG 840
Db 4896 TTCTTTGTTAAAGAACTTATTTCTCTGAAAACCTCTGCTCAGTCAATGAGG 4837
Oy 841 CCATCAGTCCACATGATCAGGCTCTCTGAGCAACATGCAATCTTTCTTTGGGAAT 900
Db 4836 CCATCAGTCCACATGATCAGGCTCTCTGAGCAACATGCAATCTTTCTTTGGGAAT 4777
Oy 901 TGAATCTTAAGCTGAATAGCTGAAGTTCAAAAAAGCTGTTGAATCTGAACCTTAC 960
Db 4776 TGAATCTTAAGCTGAATAGCTGAAGTTCAAAAAAGCTGTTGAATCTGAACCTTAC 4717
Oy 961 AGTGGCTTGCAGAGTCACTGCTCCATTTCTTAATTTCTTAAGCTGATTAATTAAT 1020
Db 4716 AGTGGCTTGCAGAGTCACTGCTCCATTTCTTAATTTCTTAAGCTGATTAATTAAT 4657
Oy 1021 CTTGTTACAGCCCTTCTGAGATGTGTGTTTTTTCACACTGCTCTTAATAGTCTGT 1080
Db 4656 CTTGTTACAGCCCTTCTGAGATGTGTGTTTTTTCACACTGCTCTTAATAGTCTGT 4597
Oy 1081 GAATTTCAATTTCTTTTATACATTTTCAATGTTTGTGTTGTGTTGTGTTT 1140
Db 4596 GAATTTCAATTTCTTTTATACATTTTCAATGTTTGTGTTGTGTTGTGTTT 4537
Oy 1141 GACTTTAGTGGCAGATCACTTCTGTTGTTTATACCAAGGAATCTGATTTGATTA 1200
Db 4536 GACTTTAGTGGCAGATCACTTCTGTTGTTTATACCAAGGAATCTGATTTGATTA 4477
Oy 1201 TCTTCCCTTTAAATTAATATCTAAGCTCAAGAGATAGGCTACCTGCTGAGG 1260
Db 4476 TCTTCCCTTTAAATTAATATCTAAGCTCAAGAGATAGGCTACCTGCTGAGG 4417
Oy 1261 TCTGAGATGATTAATCAAGAGCTGCTAATCCCAAGTTGACCAAGCTTTAC 1320
Db 4416 TCTGAGATGATTAATCAAGAGCTGCTAATCCCAAGTTGACCAAGCTTTAC 4357
Oy 1321 AACATTTGCTCTCTCTTCCAGATTAATGAGCTTGAAGAAAGAACTTAATTTGC 1380
Db 4356 AACATTTGCTCTCTCTTCCAGATTAATGAGCTTGAAGAAAGAACTTAATTTGC 4297
Oy 1381 CAAGAGCTCAGAGAGACCAAGAACTGCTGAGATATGATGATTAATTTAA 1440
Db 4296 CAAGAGCTCAGAGAGACCAAGAACTGCTGAGATATGATGATTAATTTAA 4237
Oy 1441 AGTTTGTGACCTTCTGCAATTTGTGTATCTAAGCAAGATTTCAATTTAATGCT 1500
Db 4236 AGTTTGTGACCTTCTGCAATTTGTGTATCTAAGCAAGATTTCAATTTAATGCT 4177
Oy 1501 AAACAAACAAGAGATCAGAGTTGAGAGCTATTAATAGATTAATCTGATAC 1560
Db 4176 AAACAAACAAGAGATCAGAGTTGAGAGCTATTAATAGATTAATCTGATAC 4117
Oy 1561 AGTTGACAACTGAATTAATCTCAAGAACCAATGAGATCAATGAGCAAAAAATCCAGAC 1620
Db 4116 AGTTGACAACTGAATTAATCTCAAGAACCAATGAGATCAATGAGCAAAAAATCCAGAC 4057
Oy 1621 ACAGAAAGTAATGTAACATATGCTGATTAATGATTAATCTAATTTAATTAAGTAAGA 1740
Db 4056 ACAGAAAGTAATGTAACATATGCTGATTAATGATTAATCTAATTTAATTAAGTAAGA 3997
Oy 1741 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
Db 3996 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3877
Oy 1801 CCGCAATATTAATTAATTTCTTTAATTAAGAAAGTTCTAATTTGAAAAACAATGTTCTAG 1860

```



Db	3876	GGTCCAAATATGAAATATTTCCCTTTATACGAAAGTCTATTTGAAAACAAATGTTCTAG	381.77
Qy	1861	AAAATATACACATATCTATTAATAAAACAAAAAGCAGTGTATGTCTAAGGCCAGGGG	192.72
Db	3816	AAAATATACACATATCTATTAATAAAACAAAAAGCAGTGTATGTCTAAGGCCAGGGG	375.75
Qy	1921	TGAGGGAGATGCATATGCACAAAGTGTATGAGAAAGTTTGGGGTAAATAGCGTTGTGCA	198.87
Db	3756	TGAGGGAGATGCATATGCACAAAGTGTATGAGAAAGTTTGGGGTAAATAGCGTTGTGCA	369.77
Qy	1981	ATCTTAATGGCGATGAAAGCTACCTGGTGTCTAATATGTCTACCTCTCAGACTGAACACT	204.00
Db	3696	ATCTTAATGGCGATGAAAGCTACCTGGTGTCTAATATGTCTACCTCTCAGACTGAACACT	363.73
Qy	2041	TGCAATTTGGCGAAATTCATTGTATGTAAATATACCTCATTAAGTAACTCTTAAGAGTCA	210.07
Db	3636	TGCAATTTGGCGAAATTCATTGTATGTAAATATACCTCATTAAGTAACTCTTAAGAGTCA	357.78
Qy	2101	AGCTTTTGTGCAAAATATTTTAAATCAGTTGCAATATATATAGAGATGATTTTGCA	216.66
Db	3576	AGCTTTTGTGCAAAATATTTTAAATCAGTTGCAATATATATAGAGATGATTTTGCA	351.77
Qy	2161	AATACATAAACATGTATATCCATTAAGTGTGCAATATTTTGTACGCTCTGAAACAC	222.20
Db	3516	AATACATTAACATGTATATCCATTAAGTGTGCAATATTTTGTACGCTCTGAAACAC	345.77
Qy	2221	AGAGATGAATTAAGATAGAGACGCTGCTCCATCAGCTGTTCACATCCAGTACAGAGATG	228.87
Db	3456	AGAGATGAATTAAGATAGAGACGCTGCTCCATCAGCTGTTCACATCCAGTACAGAGATG	339.77
Qy	2281	AGCTATATCAAAATAGCTAGACCTCCAGAGAAAGTTATAGGTAGACTTACACAAAATA	234.00
Db	3396	AGCTATATCAAAATAGCTAGACCTCCAGAGAAAGTTATAGGTAGACTTACACAAAATA	333.77
Qy	2341	GTCACAGATTAATTAATGTAGAGACGTATGAAAGTGGGAAAGTTCTTTTATGTGAAAAA	240.00
Db	3336	GTCACAGATTAATTAATGTAGAGACGTATGAAAGTGGGAAAGTTCTTTTATGTGAAAAA	327.77
Qy	2401	GAGCGAATTTTGTGCTTTTGAAGATGACAGATGTGAATATGCGACATTCACATTT	246.00
Db	3276	GAGCGAATTTTGTGCTTTTGAAGATGACAGATGTGAATATGCGACATTCACATTT	321.77
Qy	2461	TAAATACATTCCTGTGTGAGGCGACAAATGTATCCAAAGCGACAAAGACAAACGAAATA	252.20
Db	3216	TAAATACATTCCTGTGTGAGGCGACAAATGTATCCAAAGCGACAAAGACAAACGAAATA	315.77
Qy	2521	TGCAACCTTAGAGAAAGTGCAATGAAAGGGAGCACTGTAAATATATTTTCAATGAATGA	258.80
Db	3156	TGCAACCTTAGAGAAAGTGCAATGAAAGGGAGCACTGTAAATATATTTTCAATGAATGA	309.77
Qy	2581	GTGAGAGAAATTTGTATCATAGACACTGAGTTTGGAGATGCAATGTTCTTGAGCTCTTA	264.40
Db	3096	GTGAGAGAAATTTGTATCATAGACACTGAGTTTGGAGATGCAATGTTCTTGAGCTCTTA	303.77
Qy	2641	GGAGTCMAAGAAACAAAGTGTCCCTTCTCTACGTTATGCTCAGTGTCCAAATCCAA	270.00
Db	3036	GGAGTCMAAGAAACAAAGTGTCCCTTCTCTACGTTATGCTCAGTGTCCAAATCCAA	297.77
Qy	2701	AACACCTTCTCTCTTAAGTAATTTCTCTCTCCCTCCATACAAATCTTAAGTCTTACA	276.00
Db	2976	AACACCTTCTCTCTTAAGTAATTTCTCTCTCCCTCCATACAAATCTTAAGTCTTACA	291.17
Qy	2761	AACATCATTAACAGCGAGGTATGTGTCAAGAAAGCAATGTCTTCTTACACTTCTAT	282.20
Db	2916	AACATCATTAACAGCGAGGTATGTGTCAAGAAAGCAATGTCTTCTTACACTTCTAT	285.77
Qy	2821	GTACGTTATATATATACAAATTCCTGCTTAAAGACCTAAAGCTGTGAAAAAGTTCCAC	288.80
Db	2856	GTACGTTATATATATACAAATTCCTGCTTAAAGACCTAAAGCTGTGAAAAAGTTCCAC	279.77
Qy	2881	CTTGCAATCAAAAGATATAATTAATGAAATTTGTATGTAACCTTAAGTCCCTTAAGAGAT	294.40

Dh	2796	CTGCACATCAAGAGATATATTCAGCATTTGTATAGTAACTTATGTCCCTTAAGAGAT	2737
Qy	2941	AGGATGAAGTAAATTAATTAACAGTAATATGGAATTTATATATATGATGGCACTTAT	3000
Dh	2736	AAAGATGAACTATTAATATTAAGAAATATATATGATATTTATATATATGATGGCACTTAT	2677
Qy	3001	TTTCACTTGAATGCTGATATGTTGCAATCTAATGATGTTCTGTTGAATTCAGAGATTTG	3060
Dh	2676	TTTCACTTGAATGCTGATATGTTGCAATCTAATGATGTTCTGTTGAATTCAGAGATTTG	2617
Qy	3061	CGTCTTTTTCCTGGGTCAACTCTGCGCAATTAATTTCCATATATGCAATAGAGCCAACTCT	3120
Dh	2616	CGTCTTTTTCCTGGGTCAACTCTGCGCAATTAATTTCCATATATGCAATATAGAGCCAACTCT	2557
Qy	3121	TTTCAATATTAATCTATTTTAAATTTGTTGCAATTAATTTCTGTTCTGTTAGCTTATGTA	3180
Dh	2556	TTTCAATATTAATCTATTTTAAATTTGTTGCAATTAATTTCTGTTCTGTTAGCTTATGTA	2497
Qy	3181	ACTTTAGGATTTTAAATTAACAACTATGAAATCATGACATACGTTTAAATGATATTAAT	3240
Dh	2496	ACTTTAGGATTTTAAATTAACAACTATGAAATCATGACATACGTTTAAATGATATTAAT	2437
Qy	3241	TAAATACGTTAGGATTAATTAACCTTTTAAATTTTAAATTTTAAATTAAGTATGATGGCT	3300
Dh	2436	TAAATACGTTAGGATTAATTAACCTTTTAAATTTTAAATTTTAAATTAAGTATGATGGCT	2277
Qy	3301	CAGACCTGTAATCCCAACACTTTGGGAAAGCCGGGTGGGAGATATGCTTGAATCCACAG	3360
Dh	2376	CAGACCTGTAATCCCAACACTTTGGGAAAGCCGGGTGGGAGATATGCTTGAATCCACAG	2317
Qy	3361	TTTGAACAAGTCAAGGCGCAACAGCAAGACCCCATATCTTAAAAAACAAACAAACAA	3420
Dh	2316	TTTGAACAAGTCAAGGCGCAACAGCAAGACCCCATATCTTAAAAAACAAACAAACAA	2257
Qy	3421	AATTAACCTGGATATGTTGTGCTCACTGTATGCTCAAGCTACACAGAGGCTGAGGAGA	3480
Dh	2256	AATTAACCTGGATATGTTGTGCTCACTGTATGCTCAAGCTACACAGAGGCTGAGGAGA	2197
Qy	3481	AGGATCACTTGAAGCCCAAGGAGGTTTGAAGGCTGACAGTATCAATGAAGCGGCTGTACCTC	3540
Dh	2196	AGGATCACTTGAAGCCCAAGGAGGTTTGAAGGCTGACAGTATCAATGAAGCGGCTGTACCTC	2137
Qy	3541	ACTCTGGGTGAAGTGCAGAGAGGCTGTCTCAAAAATATTAATTAATTAATTAATTAATTAAT	3600
Dh	2136	ACTCTGGGTGAAGTGCAGAGAGGCTGTCTCAAAAATATTAATTAATTAATTAATTAATTAAT	2077
Qy	3601	AAAAAACAAATTAATTAATTTTAAATTTTAAATTAACACACACTAGAGATGTTTGCAAATTA	3660
Dh	2076	AAAAAACAAATTAATTAATTTTAAATTTTAAATTAACACACACTAGAGATGTTTGCAAATTA	2017
Qy	3661	TATATTGGAGATCTATATCCCTGGAAGTTATTTAAATTAATTTTGAAGAGTTCTTCCCTCA	3720
Dh	2016	TATATTGGAGATCTATATCCCTGGAAGTTATTTAAATTAATTTTGAAGAGTTCTTCCCTCA	1957
Qy	3721	TTTCTTAAGAGATGCGCAATTTGTAAATTTCAAGGCTTGAAGAGACATGAGGCTCGCCACT	3780
Dh	1956	TTTCTTAAGAGATGCGCAATTTGTAAATTTCAAGGCTTGAAGAGACATGAGGCTCGCCACT	1897
Qy	3781	CGAAAGTGTGCTCCAAAGACCAAGACATCAAGTAACTCTGGGAAAGTTTGAATGCAAG	3840
Dh	1896	CGAAAGTGTGCTCCAAAGACCAAGACATCAAGTAACTCTGGGAAAGTTTGAATGCAAG	1837
Qy	3841	AGCTTAAAGCTCAAGCCCAAGACCTATGAAACAGAAATGCTGATTAACAAATTTCTAGAT	3900
Dh	1836	AGCTTAAAGCTCAAGCCCAAGACCTATGAAACAGAAATGCTGATTAACAAATTTCTAGAT	1777
Qy	3901	GGCTTAAGGGGACATTTAAATCTTGAGAGCTGTGCACTAGAAATCTTCACTCCACCTTTC	3960
Dh	1776	GGCTTAAGGGGACATTTAAATCTTGAGAGACCTTGTGCACTAGAAATCTTCACTCCACCTTTC	1717
Qy	3961	ATTATTAATGCAATCACTTGGGCTGTGGTCAACAGAAATATGATATTTTAAATTTTCAGAA	4020
Dh	1716	ATTATTAATGCAATCACTTGGGCTGTGGTCAACAGAAATATGATATTTTAAATTTTCAGAA	1657

BEST AVAILABLE COPY

QY 4021 CCTTCTATTTAGTCATCTATATTTGCTAATAGCAGGAGAAAGCCAACTCTTTACT 4080  
DB 1656 CCTTCTATTTAGTCATCTATATTTGCTAATAGCAGGAGAAAGCCAACTCTTTACT 1597  
QY 4081 GCATTTAAACAATCTATATTAATTAAGTAAAGCAATCTTCCCTTTAGCTTTTACATTTTG 4140  
DB 1596 GCATTTAAACAATCTATATTAATTAAGTAAAGCAATCTTCCCTTTAGCTTTTACATTTTG 1537  
QY 4141 TGGAGCAGCTGTTGTAATTTGGCTGGGGCTCAGCGCGCTGTTGTAATTTTCAAACT 4200  
DB 1536 TGGAGCAGCTGTTGTAATTTGGCTGGGGCTCAGCGCGCTGTTGTAATTTTCAAACT 1477  
QY 4201 CACAGATTTAGCCCTCTGCGGCTAAATTAAGCAAGAAATGTCATTTTAAATAGCT 4260  
DB 1476 CACAGATTTAGCCCTCTGCGGCTAAATTAAGCAAGAAATGTCATTTTAAATAGCT 1417  
QY 4261 TCTCCCTTCAATCTGCTGAGCAAGCAATTAATTTTATGAAAACATTTTGAAT 4320  
DB 1416 TCTCCCTTCAATCTGCTGAGCAAGCAATTAATTTTATGAAAACATTTTGAAT 1357  
QY 4321 TAGATTTACTTACAGGAAATATGCAATTTCTGAAAGGGCTTTAGATGTCACAAAC 4380  
DB 1356 TAGATTTACTTACAGGAAATATGCAATTTCTGAAAGGGCTTTAGATGTCACAAAC 1297  
QY 4381 TTTGACATCTACTGATGTCACCTATTTTACAGGTGTCTGTAAGCTGAGGAGTAAAGGA 4440  
DB 1296 TTTGACATCTACTGATGTCACCTATTTTACAGGTGTCTGTAAGCTGAGGAGTAAAGGA 1237  
QY 4441 AGATGTGAATCAGCAATTTAGTACAGGTAAATACAGAGTGTGTTTTTCCCTCTG 4500  
DB 1236 AGATGTGAATCAGCAATTTAGTACAGGTAAATACAGAGTGTGTTTTTCCCTCTG 1177  
QY 4501 TTGGATCTATCTTACCTAGCTCTTCAATCAATTTTCAATTTTCCAAATCCAGAA 4560  
DB 1176 TTGGATCTATCTTACCTAGCTCTTCAATCAATTTTCCAAATCCAGAA 1117  
QY 4561 ACCAGATTAAGTTTACAGCCATTTTCAGAAAGAAATTAATTTTGTGTAGACTT 4620  
DB 1116 ACCAGATTAAGTTTACAGCCATTTTCAGAAAGAAATTAATTTTGTGTAGACTT 1057  
QY 4621 TCCGATTAATTTACAGATTTTGGAAATATATGAACAATTTTATGTTTCTTTGCAAGTAG 4680  
DB 1056 TCCGATTAATTTACAGATTTTGGAAATATATGAACAATTTTATGTTTCTTTGCAAGTAG 997  
QY 4681 GTCAAGTCAAGCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAAC 4740  
DB 996 GTCAAGTCAAGCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAAC 937  
QY 4741 CCGACTGAGAGATTAATTAATTAATTAATTTTATTAACAGCAATTTGAAATATTT 4800  
DB 936 CCGACTGAGAGATTAATTAATTAATTAATTTTATTAACAGCAATTTGAAATATTT 877  
QY 4801 GTGACCTGCAATTTCTGCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 4860  
DB 876 GTGACCTGCAATTTCTGCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 817  
QY 4861 TTAATTAATTAATTAATTAATTAATTAATTTTATTAACAGCAATTTGAAATATTT 4920  
DB 816 TTAATTAATTAATTAATTAATTAATTAATTTTATTAACAGCAATTTGAAATATTT 757  
QY 4921 AGGTATTTTCTTTATTTCCGCTCAGCAAAAGCTTTCAAGGCAATGCAAAAGCAATTTCT 4980  
DB 756 AGGTATTTTCTTTATTTCCGCTCAGCAAAAGCTTTCAAGGCAATGCAAAAGCAATTTCT 697  
QY 4981 CTTTCTCTAGCTGAGAGAAAGTGAATGCTTAAATTAATTAATTAATTAATTTTCCCT 5040  
DB 696 CTTTCTCTAGCTGAGAGAAAGTGAATGCTTAAATTAATTAATTAATTTTCCCT 637  
QY 5041 GTCTTCTCTATTTGAGATTTGCAACAACAGCGGCTTGGCTGAAAGGCAATGCAAGG 5100  
DB 636 GTCTTCTCTATTTGAGATTTGCAACAACAGCGGCTTGGCTGAAAGGCAATGCAAGG 577

QY 5101 CCGGAGGAGGAGAAATAGATGAAACCAAAACCAAAACCAAAACCTTCCCTTAAGCTCT 5160  
DB 576 CCGGAGGAGGAGAAATAGATGAAACCAAAACCAAAACCAAAACCTTCCCTTAAGCTCT 517  
QY 5161 ACATAACATTTTAGCCCGAGAAATAGTACAGAAATCTCAATTAATTAATTAATTAATTAAT 5220  
DB 516 ACATAACATTTTAGCCCGAGAAATAGTACAGAAATCTCAATTAATTAATTAATTAATTAAT 457  
QY 5221 TACAGAAAGTGTATGATGATGAGCAAGGCTGAGCACTCATGCTCAGTTCAATTACA 5280  
DB 456 TACAGAAAGTGTATGATGATGAGCAAGGCTGAGCACTCATGCTCAGTTCAATTACA 397  
QY 5281 AAAGTCAGGCTGCTGAAATTAATTAATTTGATGATGATGATGATGATGATGATGATGAT 5340  
DB 396 AAAGTCAGGCTGCTGAAATTAATTAATTTGATGATGATGATGATGATGATGATGATGAT 337  
QY 5341 GAGATCAAGATTCAGAGATGCTTCCAGCTTCAATTTGCAATTTGCAATTTGCAATTTGCA 5400  
DB 336 GAGATCAAGATTCAGAGATGCTTCCAGCTTCAATTTGCAATTTGCAATTTGCAATTTGCA 277  
QY 5401 TACGTCAGAGATCTGAGACCTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5460  
DB 276 TACGTCAGAGATCTGAGACCTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 217  
QY 5461 CTTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5520  
DB 216 CTTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 157  
QY 5521 TCAAGAAAGCTGCGCTCTTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5580  
DB 156 TCAAGAAAGCTGCGCTCTTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 97  
QY 5581 TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5640  
DB 96 TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 37  
QY 5641 AGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5676  
DB 36 AGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1

RESULT 6  
BD093092/c  
LOCUS 5676 bp RNA linear PAT 27-AUG-2002  
DEFINITION Gene encoding promoter region of tumor suppressor gene p51.  
ACCESSION BD093092  
VERSION 1 GI:22638680  
KEYWORDS MO 0100818-A/4.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 5676)  
Sakai, T., Kagaya, S., Sato, T., Suenaga, Y. and Fujii, H.  
Gene encoding promoter region of tumor suppressor gene p51  
Parent: MO 0100818-A 4 04-JUN-2001;  
TOSHIYUKI SAKAI, NIPPOON KAYAKU KK, SHIEHIDE KAGAYA, TAKAMICHI SATO,  
YOSHIKAZU SUENAGA, HIDEJI FUJII  
COMMENT OS Homo sapiens (human)  
PN MO 0100818-A/4  
PD 04-JAN-2001  
PF 28-JUN-2000 MO 2000JP004261  
PR 29-JUN-1999 JP 99P 183195  
PI TOSHIYUKI SAKAI, SHIEHIDE KAGAYA, TAKAMICHI SATO, YOSHIKAZU SUENAGA,  
PI HIDEJI FUJII  
PC C12N15/10, C12K5/10, C12Q1/68, A61K48/00, A61K45/00, A61P43/CO, PC  
A61P35/00  
CC

FEATURES  
source Location/Qualifiers  
1..5676  
/organism="Homo sapiens"



BEST AVAILABLE COPY

3636 TGAATGGCGAATTCATGTAATGAAATTAACCTCAATTAAGAGTCA 3577  
QY AGTGTGTTGGAAATTTTAACTGAGTCAATCTTATGAGATGATTTTGA 2160  
Db AGTGTGTTGGAAATTTTAACTGAGTCAATCTTATGAGATGATTTTGA 3517  
QY AATACATTAACATGTAATCATCATAGGTGCAATATTTTGGCTGCTGAAAAC 2220  
Db AATACATTAACATGTAATCATCATAGGTGCAATATTTTGGCTGCTGAAAAC 3457  
QY 2221 AGAGTGAATTAAGATTAAGAGTCAATGAGTCAATGAGTCAATGAGTCA 2280  
Db 3456 AGAGTGAATTAAGATTAAGAGTCAATGAGTCAATGAGTCAATGAGTCA 3397  
QY 2281 AGTCAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 2340  
Db 3396 AGTCAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3337  
QY 2341 GTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2400  
Db 3336 GTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3277  
QY 2401 GAGGAGATTTTGTGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2460  
Db 3276 GAGGAGATTTTGTGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3217  
QY 2461 TAAACATTTCTGTGAGGAGAGATATGATCAAGGAGAGAGAGAGAGAGAG 2520  
Db 3216 TAAACATTTCTGTGAGGAGAGATATGATCAAGGAGAGAGAGAGAGAGAG 3157  
QY 2521 TCGACCTGAG 2580  
Db 3156 TCGACCTGAG 3097  
QY 2581 GTGAGAGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 2640  
Db 3096 GTGAGAGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 3037  
QY 2641 GAGTCAAG 2700  
Db 3036 GAGTCAAG 2977  
QY 2701 AACAGCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760  
Db 2976 AACAGCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2917  
QY 2761 AACATCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
Db 2916 AACATCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2857  
QY 2821 GTAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2880  
Db 2856 GTAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2797  
QY 2881 CTGAGCATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2940  
Db 2796 CTGAGCATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2737  
QY 2941 AAGAGTGAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3000  
Db 2736 AAGAGTGAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 2677  
QY 3001 TTTGACATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 3060  
Db 2676 TTTGACATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 2617  
QY 3061 CTTCTTTTCTGAGTGAAGTCTGAGTGAAGTCTGAGTGAAGTCTGAGTGAAGTCT 3120  
Db 2616 CTTCTTTTCTGAGTGAAGTCTGAGTGAAGTCTGAGTGAAGTCTGAGTGAAGTCT 2557  
QY 3121 TTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3180  
Db 2556 TTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2497

QY 3181 ACTTAGAGATTTTAAATTAACATATGAAATCATGACATGATTAATTAATTAAT 3240  
Db 2496 ACTTAGAGATTTTAAATTAACATATGAAATCATGACATGATTAATTAATTAAT 2437  
QY 3241 TAAATAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3300  
Db 2436 TAAATAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 2377  
QY 3301 CATGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3360  
Db 2376 CATGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2317  
QY 3361 TTTGAGACAGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420  
Db 2316 TTTGAGACAGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2257  
QY 3421 AATTACCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3480  
Db 2256 AATTACCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2197  
QY 3481 AGGATCACTTGAAG 3540  
Db 2196 AGGATCACTTGAAG 2137  
QY 3541 AGTCTGAGTGAAG 3600  
Db 2136 AGTCTGAGTGAAG 2077  
QY 3601 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3660  
Db 2076 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2017  
QY 3661 TTAATTTGAGAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3720  
Db 2016 TTAATTTGAGAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1957  
QY 3721 TTTCTGAGAGAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3780  
Db 1956 TTTCTGAGAGAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1897  
QY 3781 CCAAGTGTGTCAG 3840  
Db 1896 CCAAGTGTGTCAG 1837  
QY 3841 AGTCTGAGAGTCAAG 3900  
Db 1836 AGTCTGAGAGTCAAG 1777  
QY 3901 GCTCAG 3960  
Db 1776 GCTCAG 1717  
QY 3961 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4020  
Db 1716 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1657  
QY 4021 CTTCTTATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4080  
Db 1656 CTTCTTATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1597  
QY 4081 GCAATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4140  
Db 1596 GCAATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1537  
QY 4141 TCGAG 4200  
Db 1536 TCGAG 1477  
QY 4201 CACAG 4260  
Db 1476 CACAG 1417

[illegible]

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
 Rives, C., Rodey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J.,  
 Sanders, W., Saverly, G., Scherer, S., Scott, G., Shattman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajda, D.,  
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
 Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K.,  
 Valas, R., Vera, V., Villaana, D., Waldron, L., Walker, B., Wang, J.,  
 Wang, Q., Wang, S., Warren, R., Warren, R., Mel, X., White, F.,  
 Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, E., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
 Weinstock, G. and Gibbs, R.A.

Unpublished  
 2 (bases 1 to 160574)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (17-OCT-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Nov 20, 2002 this sequence version replaced gi:24080629.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Project name: KBUG  
 Center project name: KBUG  
 Center clone name: CH230-436C20

----- Summary Statistics  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 15133 bases at least Q40  
 Consensus quality: 15144 bases at least Q30  
 Consensus quality: 15425 bases at least Q20  
 Estimated insert size: 155447; sum-of-contigs estimation  
 Quality coverage: 7x in Q20 Bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 4 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 93077: contig of 93077 bp in length  
 \* 93078 93177: gap of unknown length  
 \* 93178 113865: contig of 20688 bp in length  
 \* 113866 113965: gap of unknown length  
 \* 113966 158782: contig of 44817 bp in length  
 \* 158783 158882: gap of unknown length

FEATURES  
 source 158883 160574: contig of 1692 bp in length.

misc\_feature  
 1. 160574  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-436C20"  
 complement(5637..6655)  
 /note="clone boundary"

misc\_feature  
 end sequence: B2203757"  
 complement(90372..91276)  
 /note="clone boundary"

misc\_feature  
 end sequence: B2203758"  
 93178..96221  
 /note="wgs\_end\_extension"

## ORIGIN

Query Match 14.1%; Score 843.2; DB 2; Length 160574;  
 Best Local Similarity 70.3%; Pred. No. 3.9e-151;  
 Matches 1417; Conservative 0; Mismatches 488; Indels 111; Gaps 17;

Qy	3981	GGCTGTGTGTCAGGAAATTGATTTTAAATTCAGAACCTTCAATTAAGTCATCTA	4040
Db	65379	GGCCATAGCTACAGGACATTAATTCATTTTAAAGCTTTCAGATTCGATCGT	65320
Qy	4041	TATTTGCTAATACAGGAGGAAAGCAACTCTTTAATCTGCAATTAA-CAATCTATTA	4099
Db	65319	TATTTGCT-----CAAGAGAAAGAGTCAAACTCAATTCATTAAGCAATCTATTA	65265
Qy	4100	TTAATTGCTAATACAGGAGGAAAGCAACTCTTTAATCTGCAATTAA-CAATCTATTA	4159
Db	65264	TTAATTGCTAATACAGGAGGAAAGCAACTCTTTAATCTGCAATTAA-CAATCTATTA	65205
Qy	4160	TGCTGTGCGCTCAGCGCCGCTTTTGTGAATTTGCAATTCACAGAGTTCAGCGCT	4219
Db	65204	TGCTGTGCGCTCAGCGCCGCTTTTGTGAATTTGCAATTCACAGAGTTCAGCGCT	65145
Qy	4220	CGGCTAATGAAAGGAGAAATGTCAGTTTAAATAGCTTCTCCATCTGGCT	4279
Db	65144	CAGCTGAGGAAAGGAGAAATGTCAGTTTAAATAGCTTCT-----CTGGCT	65093
Qy	4280	GAGGACAAATTAATTTTATGAAACATTTTGAAGTATTAATTAATTAATTAATTA	4339
Db	65092	AGACACAAATTAATTTTATGAAACATTTTGAAGTATTAATTAATTAATTAATTA	65034
Qy	4340	ATGTCAAATTTCTGAAAGGCTTTAGATTTGTCACACTTTGACATCTAGATGTC	4395
Db	65033	ATGTCAGATTTCT-----CTTCAAGCTTCTGTCACAACTTTGACATCTAGATGTC	64983
Qy	4400	ACCTATTACAGGTGTGTCTGTGACTAGAGGGGTGAAGGAAAGATGATCAATGTC	4459
Db	64982	ACTATTGTCAGGTGT-----CTATGACTAGAGGGGTGAAGGAAAGATGATGTC	64927
Qy	4460	TAGTACCGCTTGAATACAGAGATGTTTTCCTCCCTGTGGAGCTTATCTAATCTG	4519
Db	64926	CAGCCACATTTGGGACACAGG-----CTATTTCCTCTTGGAGCTTATCTAATCTG	64874
Qy	4520	ACCTTCTGAATCAATTTTCATTAATTTCAATTCGAATTCGAATTCGAATTCGAATTC	4579
Db	64873	ACCTTCTGAATCAATTTTCATTAATTTCAATTCGAATTCGAATTCGAATTCGAATTC	64814
Qy	4590	CCATATTCAAGAAAGAAATTAATTTTGTGTGACTTTT-CCATATTCAAGTAT	4638
Db	64813	CTACATTTCAAGAAAGAAATTAATTTTGTGTGACTTTT-CCATATTCAAGTAT	64754
Qy	4639	TTGGAAATTTTGAACAAATTTTATGTTCTTTCGAATAGTATGCAAGCAAGCAAC	4698
Db	64753	TTTCATTAATTAATTAATTTTATGTTCTTTCGAATAGTATGCAAGCAAGCAAC	64694



[illegible]

Oy	5736	TACAGTACTGCGCCTTGACCCTTAATTCAGGGGTGAATTGAATNGACATTAATTTCTGC	5795	
Dd	63642	TACAGTACTGCCCTTAGCCCTTACATTCAGGGGTGAATTGAATTCAGACTGACGCTCTGC	6358	
Oy	5796	AAAACCTAATGAAAGTGCCTTGTGATTAAGAATGTCAAGCTGTGTCAAAGAACATT	5855	
Dd	63582	AAAACCTAAGCATGGGTGTGC-TGTGTATATTCAGAGTGTTCCTGCTGTACTA--AAGAAAGC	6352	
Oy	5856	CCTCCTGTTGTTAGTCAGACAGAGATATTTATTTTGGACCTTCGTGGACTTAAGTGGTC	5915	
Dd	63525	CCCTCCTGTTTAGTGGGGCATTAATATCTACTTTGGACCTTTGTGTGGACTTAAAGTGGTT	6346	
Oy	5916	TGTGGACATATTTTCTGAAATGTCTTTTTGGTTGAT	5951	
Dd	63465	TATTAAGACATCTTTCTTACAGTAACTTATATCTTAT	63430	
RESULT #	ACI13784	284005 bp	DNA linear	HTG 19-NOV-2002
LOCUS	ACI13784/c			
DEFINITION	Rattus norvegicus clone CH230-98F2, WORKING DRAFT SEQUENCE, 3 unordered pieces.			
ACCESSION	ACI13784			
VERSION	ACI13784.5	GI:25072720		
KEYWORDS	HTG; HTGS; PHASBL; HTGS DRAFT; HTGS_FULUTOF.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Rattus norvegicus			
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
	Rattus.			
REFERENCE	1 (bases 1 to 284005)			
AUTHORS	Muzny,D.Merie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,			

Burkholderia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 284005)

Muzny, D., Marle, M., Metzker, M., Lee, S., Abman, S., Adams, C., Alder, J., Allen, C., Allen, H., Alibekov, S., Amin, A., Arguano, D., Artyushchench, V., Aoyagi, A., Ayodeji, M., Baca, B., Baden, H., Baldwin, D., Bandaru, A., Barber, M., Barnstead, M., Bernhardt, P., Biswal, K., Blair, P., Blackburn, K., Blyth, P., Brown, M., Bryant, K., Bunney, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceaas, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Crete, A., P. Souza, L., Davila, M., D., Davis, C., Day-Carroll, L., De Aranda, C., Dederich, D., Delgado, O., Denison, S., Denaro, C., Ding, Y., Dinn, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C., A. Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C., G., Gabisi, A., Genter, R., Garcia, A., Gane, I., Garra, M., Garret, G., P., Haland, N., Hamill, C., Hamilton, C., Hamilton, K., Gurevich, E., Geer, K., Gill, R., Girdy, M., Guerra, W., Guevara, W., Gutierrez, P., Haland, N., Hamill, C., Hamilton, C., Hernandez, J., Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hildun, S. L., Hodgson, A., Hogues, N., Hollins, B., Howells, S., Hu, Y., Hulse, S., Hume, J., Idlebird, D., Jackson, A., Jacob, L., Jiang, R., Johnson, B., Johnson, R., Jolivet, A., Karapetyan, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, M., Liu, Y., London, P., Longacre, S., Lopez, J., Loshenshwa, L., Lonsstedt, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindratna, M., Mahmoud, M., Malloy, K., Mangum, B., Mapus, P., Martin, K., Martin, R., Matline, E., Manthey, S., McLeod, M., P., McNelly, T. Z., Neeman, B., Milosavljevic, A., Miner, G., Ming, A., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Naik, L., Nankervis, C., Neal, D., Newton, S., Nguyen, N., Norris, S., Nwokwelench, O., Okunolu, G., Olarinmoye, A., Pal, S., Parks, K., Patel, N., Patel, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plummer, F., Polindexter, A., Popovic, C., Pritmus, E., Pu, L., Puzo, M., Quito, R., Racine, E., Reeves, K., Reiter, M., A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Rhodes, R., Riggs, F., Rivers, C., Rodkey, T., Rojals, A., Rose, M., Rose, R., Ruiz, S., J., Sanders, W., Savory, G., Scheber, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Sivaratnam, A., Sisson, I., Sitter, C. D., Smit, J., Sneed, A., Sodegren, E., Song, X., Z., Sorzelle, R., Sosa, J., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Tjebjo, Z., Uemami, K.,

**BEST AVAILABLE COPY**



TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 284005)
AUTHORS	Morley K. C.
TITLE	Direct Submission
JOURNAL	Submitted (05-MAR-2002)
REFERENCE	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
AUTHORS	3 (bases 1 to 284005)
TITLE	Rat Genome Sequencing Consortium.
JOURNAL	Direct Submission
REFERENCE	Submitted (19-NOV-2002)
AUTHORS	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

**COMMENT**

On Nov 19 2002 this sequence version replaced g1:23195097. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a contig-scaffold). Within each contig-scaffold, individual sequence contigs are ordered and oriented and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
Project Information  
Center project name: GSGN  
Center clone name: CH20-98F2

Assembly program: Phrap, version 0.990322  
Consensus quality: 249691 bases at least Q40  
Consensus quality: 252542 bases at least Q30  
Consensus quality: 254124 bases at least Q20  
Estimated insert size: 251183; sum-of-configs estimation  
Quality coverage: 7x in Q20 bases; sum-of-configs estimation

NOTE: Estimated insert size may differ from sequence length. (see [http://www.bosc.bcm.tmc.edu/doc/Gensbank\\_draft\\_data.html](http://www.bosc.bcm.tmc.edu/doc/Gensbank_draft_data.html))

NOTE: This sequence may represent more than one clone.

NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

**FEATURES**

**SOURCE**

```
misc_feature
1. 1035
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-98P2"
```

```

/note="wgs_end_extension
clone_end:T7"
misc_feature
7353..8173
/note="clone_boundary
clone_end:T7
site:
end_sequence:RH306729"
9665..11350
/note="wgs_contig"
14183..15278
/note="wgs_contig"
15329..16533
/note="wgs_contig"
complement(1261454..264092)
/note="Clone boundary
clone_end:Sp6
site:
end_sequence:RH306731"
274560..275276
/note="wgs_end_extension
clone_end:Sp6"

```

Query Match	14.1%	Score 843.2	DB 2	Length 284005
Best Local Similarity	70.3%	Pred. No. 3,7e+151		
Matches 1417	Conservative	0	Mismatches 488	Indels 111
			Gaps	17
QY	3981	GGCTGTGCTACAGGAAATTGATTTATTTTAAATTCGAAACCTTCATTTAGTCATATCA	4040	44
DB	72189	GGCCATAGCTCAAGACATTTGATTAATCTCAATTTAAGAACTTTCGATTCGATCGGCTCATCTG	72131	44
QY	4041	TATTTGCTAATATGACGGAGAGAAAGCCAACTCTTAACTGCAATTA-CAATCTATA	4099	44
DB	72129	TATTTGCT-----CAGAGAGAAAGTCAAACCTTCATTCATATGACAAATCTATA	72077	44
QY	4100	TTAATAGTTAAGCATCTCCCTTTAGATTACATTTGTGAGACAGCTGTATAT	4155	44
DB	72074	TTAATTTAGTAAATGATCTTTGTCTTCAAGCTTCACATTTTGTAGAAAAGCTCTGAT	72070	44
QY	4160	TGCTGTGGCTCAGGCGCGCCCTGTGTGAAATTTCAAATTCACATGTAGCGCT	4215	44
DB	72014	TGCTGTGGCTCAGGCGCCCATGTGTGTGAAATTTCACAAATTCACATGTAGACGCT	7199	44
QY	4220	CGGCTAATGAAAGAGAGAAATGCAATTAAATGCTTCTCCCTTCATCTGCT	4277	44
DB	71954	CAGTCTGAGAGAAAGAGAGAAATGCAATTTAAATAGCTTCTC-----CGGCT	7196	44
QY	4280	GAGCAACAAATTAATATTTTATGAAACAACTTTGATGATTATCTTACAGGAA	4335	44
DB	71502	AGAAACAACAAATTAATCATCTTTATG-AAACATTTTGAAACCAAAATTTCTTACAGGAA	7184	44
QY	4340	ATGTGCAATTTCTCTGAAAGGCTTATATTTGTCTGCACTTTGCATCTATGATGTC	4399	44
DB	71843	ATGTGCAATTTCT-----CTTCAGCTTGCTCTACAACTTTGACATTTCTGATGTC	7177	44
QY	4400	ACCTATTTACAGCTGTGCTCTGTGCTAGAGGGGTGAGAGGAGATGTGAATCACTACATGT	4455	44
DB	71792	ACTAGCTTTGCAAGGTGT---CTATGACATACGGCAATGAG-AGATATATGATCTTATGATGT	7177	44
QY	4460	TATGTAACCGTTAGATACACAGATGCTTTTTCCTCCCTGTGGAATCTATCTAAGT	4511	44
DB	71736	CAGGCACTTGGGCAACGCG-----CTCATTTTCCCTTTGGAATCTATCTATTT	7161	44
QY	4520	AGCTTCGATTCATTTTCAATTCATTTTCCAAATTCACAAACAGAGATTAAGTTACAG	4577	44
DB	71663	AGCTTCGATTCATTTTCAATTCATTTTCCAAATTCACAAACAGAGATTTAGTTACGCG	7166	44
QY	4580	CCATATTCAGAAAGAAATTAATTTATTTTGTGTGAGCTTT-CTGATTTACAGTAT	4633	44
DB	71623	CTACATTCAGAAAGAAACCAATTCATCTATCTGAGAGCTTTGCTGTATATATTAAT	7155	44
QY	4639	TTGGAAATATATGACCAATTTATGCTTTCCTTCAAGTATGCTCAATCAAGCAAAAC	4699	44

Db	71563	TCCTCCATATTAACAAATCCAGATTTTAGAGTAGAGTAATAGTCAGAGGAGCTAAGAG	71560
OY	4699	CAAAAACACAAAACCTGTAGACATATTAAGAAATAGAGTGCACGCCACTGACAGATTAAA	4758
Db	72503	CAACTAAAGTAGACGTGAGAGATAGAGATAGACACACTGATAGAGTTAGAGCTCACTTAT	71444
OY	4759	TAAACTAGAAATTTTATTTAAACGCAATTTGAAATATTTGGCACTTCAGAAATATTC	4818
Db	71443	TTTTCCATTGCGGGCTATCTGAAATTAATTTCTAAATTTAAATTCCAATATATTAACAT	71384
OY	4819	TACATATATATATTTTCCAAATTTAAATCTTTAAGAAATTAATAATATATATATGTA	4878
Db	71383	TAAAGAGCATTTTGAAGAAAGAAACATTAACCTATTAAGCTCTATTTATATACATTA	71324
OY	4879	GTACATGTCATGTTGTAGAGTAAATTTTAACCTAAATTAAGTTATTTCTTTATAT	4938
Db	71323	ATATCTATTAATCTACATCT---GTATGATATAAATACACAGGCTATTTTTTTATAT	71267
OY	4939	CGGCTAGGCAAGCTTCTTAGAGGAATGCTAAAGGAATCTCTCTCTCT-----	4988
Db	71266	TGAGAGAGAAATATTTTCTAATATACATAGGCAATCTCTCTCTCTCTCTCTCTCTCT	71207
OY	4989	-----TACGTGAGAGAGAGTCAATTAAAG	5015
Db	71206	CT	71147
OY	5016	TTAATATATATCAAGAAATTTCCCTGTCTTTGTCTATTTGAGATTTGTACACACACAGGC	5075
Db	71146	TTAAATATATATCAAGAAATTTCCCTGTCTTTGTCTATTTGAGATTTGTACACACACAGGC	71087
OY	5076	GTTCGCTGAAAGGAAATCTCAAGGCGCGGAGAGGAGGAAATGATGAAATCAAAACA	5135
Db	71086	GTTCGCTGAAAGGAGGAGCTCAAGGCGCGGAGAGGAGGAAATGATGAAATCAAAACA	71036
OY	5136	AAACAAAATTCCTCCAGAGAGCTCTACAAAACCTTTAGCCCGCAAAATAGTCAACGAA	5195
Db	72103	AAAAATATCTCTTAGAGCTCTCTGCAACAGTCTAGCCACAGCCACATGACACAGAG	70978
OY	5196	TCTCTAAATCAACAGTATTCAGATTCAGAGATGAGGATGTTATATAGCTAGACAGGTGA	5255
Db	70975	CTCTCAATCAATATAGATATAGATATCAAGAGAGATGTTATATAGTTGCAACAGAGGGA	70916
OY	5256	CAGTATCAAGCTCAGTTCAAGTTACAAAGTCCAGGCTGCTCAAAATTTAACTGATGCCA	5315
Db	70915	CAGTATTAATCAAGCTCAGTATTAATCAAGTCCAGGCTGCTCAAAATTTAACTGATGCCA	70863
OY	5316	TTCTATGCCAGTCAATCAAGACAGAGATCAGAGATTCAGAGATGCTCAGTCCAA	5375
Db	70862	TTCTATGCCAGTCAATCAAGACAGAGATTCAGAGATTCAGAGATGCTCAGTCCAA	70808
OY	5376	TTGCCAAACAAAGTGTGGCTATATATATAGTCAAGGACTGTAAAGCTGTAAAGAGAGGGA	5435
Db	70807	----TCCCAACAAAGTGTGGCTATATATATAGTCAAGGACTGTAAAGCTGTAAAGAGAGGGA	70753
OY	5436	AGAAACAAGTAGAGAGATGCCAGCTGGTAAGATGAGGTGTTATGAGAGTTTATAGT	5495
Db	70752	AGAAACAAGTTACAGAGATGCCAGCTGGTAAGAGTCAAGTGTGTTATGAGTCCCGCT	70693
OY	5496	AATGATGATCTCATTTGGCTAAATATCAAGAAACGCTCGGCTCTTTGCAAAATATGATATG	5555
Db	70692	ACTTGAAGAGCTGATTTGGCTGAAGAGAGAAATCTCCGCTCTTGAACAAATCTGTGTA	70633
OY	5556	AGGAGAGAGGCTCAAACTTCTATCTGATAGAGATTTGAGACCTTATGCTTTAGGCT	5615
Db	70632	AAAGGAGAGAGTGTCTAAAGTCTATATGTCGATAGGCACTTATGCTTTAGGCT	70573
OY	5616	CCGAGCTTATATCTAATATATACAGAGATTTGTGATATTTTATATATATGTTCTCCG	5675
Db	70572	CCGAGCTAATATCTAATATATCTCAGAGATATATATATATATATATATATGTTCTCCCA	70513
OY	5676	TTCTGATATCAAAAGACAGTTGAGAGAAATGATTTGAACTTCAAGGTGTGCCAGCC	5735
Db	70512	TTCTGATATCAAAAGAGTGAAGAAATGATTTGAACTTCAAGGTGTGTCTACCC	70453

Oy	5730	TACAGTACGGCCCTGACCCCTTAACATCAGCGGGATTTGAATGTGACATTAACCTTCTC	5795
Db	70452	TACAGTACGCGCCCTGACCCCTTAACATCAGCGGGATTTGAATGACATTAACCTTCTC	70393
Oy	5796	AAACTATTATGAAGTGCCTTGATATTAATGATATGTCAGCTGTGTACCAAGAACAT	5855
Db	70392	AAACTGAACGCTGTGTGC--TGTGATATGAGATGTTCTGTCTTACAT--AAGAAC	70338
Oy	5856	CCTCCTGTTTAGTCAGACAGATGATATATTTGACCTTCTGTGACCTTAAGTGTG	5915
Db	70335	CCTCCTGTTTAGTGGGCAATATGATATTAACCTTGGACCTTGTGTGACCTTAAGTGTG	70276
Oy	5916	TGTGACATATTTCTGAATGTCTTTTTCCTGAT	5951
Db	70275	TATGACATCTTTCTACAGTAACTTAATCTTAA	70240

RESULT 9	AC126304/c	LOCUS	DEFINITION	ACCESSION	KEYWORDS	SOURCE	ORGANISM
	AC126304	20413 bp	DNA	linear	HTG 20-NOV-2002		
	Rattus norvegicus clone	Ch230-244D17	***	SEQUENCING IN PROGRESS			
	***, 3 unordered pieces.						
	AC126304						
	AC126304.6	GI:25118584					
	HTG: HTGS_PHASE1	HTGS_DRAFT	HTGS_ENRICHED.				
	Rattus norvegicus	[Norway rat]					
	Rattus norvegicus						

## REFERENCE AUTHORS

REFERENCE  
AUTHORS

(Pages 1 to 204313)

Muzny, D., Matle, Metker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Anguitano, D., Ayala-Becchi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Bernhardt, F., Biewald, K., Blatz, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bubay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., Dedeich, D., Davila, L. D., Davis, C., Davy-Carroll, L., De Anda, C., Dedertich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Din, H., Dlyva K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duvall, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falset, T., Fan, G., Fernandez, S., Finley, M., Flagg, S., Forbes, L., Foster, T., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gnatelogeys, E., Geier, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howell, S., Huily, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, K., Kows, C., Kraft, C.L., Lebrow, H., Lewan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzueva, L., Lorisged, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangun, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minda, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Muniasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackelmech, O., Okunonu, G., Olanunnsogun, A., Pal, S., Parke, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Polndexter, A., Popovic, D., Pritms, R., Pu, L. L., Puazo, M., Quizon, J., Rachlin, F., Reeves, K., Reiter, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savory, G., Scheerer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvatsbeyn, A., Sison, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabot, F., Taylor, C.,

BEST AVAILABLE COPY



4759 TAACTAGATATTTTATTAACAGCAATTGAAATTAATTTGCACTTCAGAAATATTC 4818  
 Db TTTTCCATTGACGGGCTACTGAAATATTTCTAAATTAATTCACATTAATAAACAAT 63146  
 Qy 4819 TACAAATATATTAATTTCCAAATTTTAATATCTTTAGAAAATTAATATATATGTA 4878  
 Db 63145 TAAAGACATTAATGAAGAAACATTTAAGTAATACCTTAATTAATATCAATTA 63086  
 Qy 4879 GTACATATGATGTTTGAAGAGATTTTAATCAATTAAGATTAATTTTAT 4938  
 Db 63085 ATACTATATTAATCACT---GTATGATTAATAATACAGAGCTATTTTAT 63029  
 Qy 4939 CGGCTGAGGAAAGCTTTAGAGGATTTGAAGGATATCTCTTCTC----- 4988  
 Db 63028 TGGAGAGAAATATTTCTAAATACATAGCACTCTCTCTCTCTCTCTCTCT 62969  
 Qy 4989 -----TACTGAGAGAGAGTCACTTCTAG 5015  
 Db 62968 CT 62909  
 Qy 5016 TTAATATATCAAGAAATTTCCCTGCTTTGCTATTTGAGATGTCACCAACAGCGG 5075  
 Db 62908 TTAATATATCAAGAAATTTCCCTGCTTTGCTATTTGAGATGTCACCAACAGCGG 62849  
 Qy 5076 GTTGGCTGAAAGGAACTGAAGGCGGAGGAGGAAATACATCAAAAACAAACA 5135  
 Db 62848 GTTGGCTGAAAGGAGGCTGAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 62798  
 Qy 5136 AAACAAATCTTCCCTAGAGCTTACAAACATTTAGCCCAATATAGTACAGAA 5195  
 Db 62797 AAAAATATCTCTCTAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 62738  
 Qy 5196 TCCCAATCAAAACAGATTCAGATCAAGAGAACTGATGTCAGAGAGAGGCTGA 5255  
 Db 62737 CTCCTAATCAATCAAGATTCAGATCAAGAGAACTGATGTCAGAGAGAGGCTGA 62678  
 Qy 5256 CACTCATAGCTCAGTCACTCAAAAGCTCAGGCTGTAATTAATCTGATGCA 5315  
 Db 62677 CATTATTAATCACTG-----GACAACTCTGAGCTCCCTGATTAATCTGATGCA 62625  
 Qy 5316 TTTATGCGAGCATCATCAAGAGAGATCAAGAGATCAAGAGATGCTCAGCTCA 5375  
 Db 62624 TTTATGCGAGCATCATCAAGAGAGATCAAGAGATCAAGAGATGCTCAGCTCA 62570  
 Qy 5376 TTTGCAACCAAGATGCTCAATCAAGAGATCAAGAGATCAAGAGATGCTCAGCTCA 5435  
 Db 62569 -----TGCACAAAGATGCTCAATCAAGAGATCAAGAGATGCTCAGCTCA 62515  
 Qy 5436 AGAACAACGTAAGAGATGCTCAAGAGATCAAGAGATGCTCAGCTCA 5495  
 Db 62514 AGAACAACGTAAGAGATGCTCAAGAGATCAAGAGATGCTCAGCTCA 62455  
 Qy 5496 AATTGATGATCTCATTTGCTAAATCAAGAACTGCTCCTTTGCAATATATG 5555  
 Db 62454 ACTTGAAGAGCTGATTTGCTGTAAGAGAAATGCTCCTCTTGAACAACTG 62395  
 Qy 5556 AAGAGAGAGATGCTCAATTTTATATGCTGATGATGATGATGATGATGATG 5615  
 Db 62394 AAGAGAGAGATGCTCAATTTTATATGCTGATGATGATGATGATGATGATG 62335  
 Qy 5616 CCCGCTTATATCTATATATACAGAGATTTGATATTTATATTAATTTGCTCG 5675  
 Db 62334 CCGCTTATATCTATATATACAGAGATTTGATATTTATATTAATTTGCTCG 62275  
 Qy 5676 TTTGTTGATATCAAGAGATGTAAGAGAAATTTTGAATTTGAGAGTGTGAGAG 5735  
 Db 62274 TTTGTTGATATCAAGAGATGTAAGAGAAATTTTGAATTTGAGAGTGTGAGAG 62215  
 Qy 5736 TTAGATATGCTGAGAGCTTATATATGAGAGAGTGTGAGAGTGTGAGAGTGTG 5795  
 Db 62214 TTAGATATGCTGAGAGCTTATATATGAGAGAGTGTGAGAGTGTGAGAGTGTG 62155  
 Qy 5796 AAAAATTAATGAAGTCTTGTATATATGATGATGATGATGATGATGATGAT 5855

Db 62154 ABAACCTAGACTGCTGCTCC-TGATATTTCTAGATTTTCTCTCTTACTA--AAGAGC 62098  
 Qy 5856 CCTCTGTTTATGTCAGACAGATGATATATTTTGAATTTCTGTGACCTTAAGTGTG 5915  
 Db 62097 CCTCTGTTTATGTTAGGAGCAATATGATTAATTTGATGATGATGATGATGATG 62038  
 Qy 5916 TGTGACATATTTTCTGAATGCTTTTGTGAT 5951  
 Db 62037 TATGACATCTTTTCTAGATTAATCTTATCTTAT 62002

RESULT 10  
 LOCUS AC113020 214853 bp DNA linear ROD 04-OCT-2003  
 DEFINITION Mus musculus chromosome 16, clone RP23-186N8, complete sequence.  
 AC113020  
 VERSION AC113020.9 GI:37515087  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 214853)  
 Birren, B., Nussbaum, C. and Lander, E.  
 Mus musculus chromosome 16, clone RP23-186N8  
 Unpublished  
 2 (bases 1 to 214853)  
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouhgalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J., Chararo, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardy, S., Girde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamet, A., Karatas, A., Kelle, C., Levine, R., Liu, G., Maclean, C., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McKernan, K., McKernan, K., Melder, J., Meneu, L., Mihova, T., Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhang, P., Pierre, N., Pollara, V., Raymond, C., Retter, R., Rickback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnapack, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomson, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Teelie, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, R., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zaimoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (25-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 214853)  
 Birren, B., Nussbaum, C. and Lander, E.  
 Direct Submission  
 Submitted (05-SEP-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 214853)  
 Birren, B., Nussbaum, C., Lander, E., Abouelella, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J., Dodge, S., Dooley, K., Dorris, L., Galagan, J., Gardy, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardy, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamet, A., Karatas, A., Kelle, C., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Melder, J., Meneu, L., Mihova, T., Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhang, P., Pierre, N.,

TITLE  
JOURNAL  
COMMENT

Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, U., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M., Talarmin, J., Teste, S., Theodore, J., Topfman, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zairoun, J., Zembek, L., Zimmer, A. and Zody, M. Direct Submission  
Submitted (04-OCT-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Oct 4, 2003 this sequence version replaced gi:34482079.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Genome Center  
Center code: WBH  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
Project Information  
Center project name: L23503  
Center clone name: 186\_N\_8

## FEATURES

```

source
1. 214853 Location/Qualifiers
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /chromosome="16"
    /map="16"
    /clone="RP23-186N8"
    /clone_lib="RPC1-23 Female Mouse BAC"
    1. 51
    /note="xgs and extension
    clone end:5p6"
    52. 57
    /note="clone boundary
    clone end:5p6
    site:ECORI"
    1042. 1080
    /rpt_family="(CATATA)n"
    1179. 1367
    /rpt_family="B3"
    2436. 2578
    /rpt_family="LM4C"
    2651. 2737
    /rpt_family="LM4C"
    complement(2850..3351)
    /rpt_family="RMKRI9B"
    complement(3350..3516)
    /rpt_family="B3"
    3527. 3891
    /rpt_family="LM4C"
    4170. 4563
    /rpt_family="LM4C"
    4873. 4952
    /rpt_family="B1_MM"
    4968. 5018
    /rpt_family="(TGA)n"
    5056. 5415
    /rpt_family="MTD"
    5554. 5810
    /rpt_family="LM4C"
    6353. 6500
    complement(6731..7895)
    /rpt_family="Lx8"
    7896. 7919
    /rpt_family="Lx7"
    /rpt_family="(TTTTC)n"
    complement(7920..8304)
    /rpt_family="Lx7"
    complement(8305..9900)
    /rpt_family="Lx4"
    complement(9899..10500)

misc_feature
52. 57
    /note="xgs and extension
    clone end:5p6"

repeat_region
    /rpt_family="(CATATA)n"
repeat_region
    /rpt_family="B3"
repeat_region
    /rpt_family="LM4C"
repeat_region
    /rpt_family="LM4C"
    complement(2850..3351)
repeat_region
    /rpt_family="RMKRI9B"
    complement(3350..3516)
repeat_region
    /rpt_family="B3"
repeat_region
    /rpt_family="LM4C"
    4170. 4563
repeat_region
    /rpt_family="LM4C"
    4873. 4952
repeat_region
    /rpt_family="B1_MM"
    4968. 5018
repeat_region
    /rpt_family="(TGA)n"
    5056. 5415
repeat_region
    /rpt_family="MTD"
    5554. 5810
repeat_region
    /rpt_family="LM4C"
    6353. 6500
repeat_region
    complement(6731..7895)
    /rpt_family="Lx8"
repeat_region
    7896. 7919
    /rpt_family="Lx7"
repeat_region
    /rpt_family="(TTTTC)n"
    complement(7920..8304)
    /rpt_family="Lx7"
repeat_region
    complement(8305..9900)
    /rpt_family="Lx4"
repeat_region
    complement(9899..10500)

repeat_region
    /rpt_family="Lx4"
    10516..10639
repeat_region
    /rpt_family="L1"
    10634..11419
repeat_region
    /rpt_family="L1"
    11420..11615
repeat_region
    /rpt_family="B2_Mm2"
    11616..13898
    /rpt_family="L1"
    complement(13899..14142)
    /rpt_family="LTRIS_MM"
    14447..17150
    /rpt_family="Lx8"
    17251..17303
    /rpt_family="(TGC)n"
    17523..17580
    /rpt_family="AT_rich"
    17590..17667
    /rpt_family="(GAAA)n"
    17669..18434
    /rpt_family="Lx8"
    18558..18593
    /rpt_family="AT_rich"
    18855..19268
    /rpt_family="BGLIT"
    21016..21156
    /rpt_family="B1_MM"
    21157..21199
    /rpt_family="(CAA)n"
    complement(21371..21767)
    /rpt_family="LM6"
    complement(21768..21878)
    /rpt_family="L1"
    21869..22524
    /rpt_family="L1_MM"
    22533..22555
    /rpt_family="(A)n"
    complement(22556..23200)
    /rpt_family="L1"
    complement(25514..25708)
    /rpt_family="B3"
    26244..26293
    /rpt_family="MT2B"
    26294..26369
    /rpt_family="B1_MM"
    26395..26369
    /rpt_family="ID_B1"
    26370..26532
    /rpt_family="(GAAA)n"
    26533..26553
    /rpt_family="ID_B1"
    26674..26921
    /note="single clone coverage"
    26678..26712
    /rpt_family="(TTTC)n"
    complement(26715..26836)
    /rpt_family="B1_MM"
    26898..26945
    /rpt_family="(TC)n"
    26945..26989
    /rpt_family="(TC)n"
    complement(26992..27264)
    /rpt_family="B4A"
    30432..30458
    /rpt_family="(T)n"
    30639..30668
    /rpt_family="(CCCCCT)n"
    complement(31137..31532)

Query Match 14.64; Score 835.4; DB 10; Length 214853;
Best Local Similarity 70.24; Pred. No. 1.2e-149;
Matches 1439; Conservative 0; Mismatches 471; Indels 141; Gaps 18;

```

QY 3981 GCTGTGTCAGAGAAATGATTATTTATTTACAGACCTCTAATTAGCATCTA 4040  
 DB 51809 GTCAATGCCATAGAGAAATGATTATTTATTTACAGACCTCTAATTAGCATCTA 51868  
 QY 4041 TATTTGCTAATAGAGAGAAAGCAAACTCTTTAATGCAATTAACAATCTAAT 4100  
 DB 51869 TATTTGCTAAGAA--GGAAGAAAGCAAACTCTTTAATGCAATTAACAATCTAAT 51923  
 QY 4101 TATTTGCTAAGCAATCTCTCTTTAAGTT--TTCATTTTGTGAGCAAGCTGTTGAT 4158  
 DB 51924 TATTTGCTAAGCAATCTCTCTTTAAGTTTATTTTATTTGAGCAAGCTGTTGAT 51983  
 QY 4159 TTGCTGTGAGTCAAGCCGCTGCTGTTGATTTTCAATTTCAATTTGAGCCGCTC 4218  
 DB 51984 TTGCTGTGAGTCAAGCCGCTGCTGTTGATTTTCAATTTCAATTTGAGCCGCTC 52043  
 QY 4219 TCGGCTTAAGTAAGAGAGAAATGCAATTTAATAGCTCTCTCTCAATCTCTGAC 4278  
 DB 52044 TCGGCTTAAGTAAGAGAGAAATGCAATTTAATAGCTCTCTCTCTCTCTCTCTCT 52095  
 QY 4279 TGAAGCAACAATTAATATTTTATTTATGAAACATTTTGAATTAATTTACTTAAGGGA 4338  
 DB 52096 TGAAGCAACAATTAATATTTTATTTATGAAACATTTTGAATTAATTTACTTAAGGGA 52154  
 QY 4339 AATGTCAATTTCTCTGAAAGGCTTTAATTTGCTCTCAACTTTGACATCTAAGAT 4398  
 DB 52155 AATGTCAATTTCTCTCTGAAAGGCTTTAATTTGCTCTCAACTTTGACATCTAAGAT 52205  
 QY 4399 CACCTATTAGAGGTGTGCTCTGATCTAGGGGGTGAAGGAGATGCACTCAACATG 4458  
 DB 52206 CACCTATTAGAGGTGTGCTCTGATCTAGGGGGTGAAGGAGATGCACTCAACATG 52261  
 QY 4459 TTAGTACCCGCTTATGATCAAGAGTGTCTTTTCTCTCTCTCTCTCTCTCTCTCT 4518  
 DB 52262 TTAGTACCCGCTTATGATCAAGAGTGTCTTTTCTCTCTCTCTCTCTCTCTCTCT 52314  
 QY 4519 GAGCTTGTGATCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 4578  
 DB 52315 GAGCTTGTGATCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 52374  
 QY 4579 CCCATTTTCAAGAAAGAAATTAATTTATTTGTTGATGATCTCTGATTTACATGAT 4638  
 DB 52375 CCCATTTTCAAGAAAGAAATTAATTTATTTGTTGATGATCTCTGATTTACATGAT 52434  
 QY 4639 TTGGGATATATGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 4698  
 DB 52435 TTGGGATATATGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 52488  
 QY 4699 CAAAAACGCAAAATCTTTAAGCATTAAGATTAAGATTAAGATTAAGATTAAGAT 4758  
 DB 52489 CAAAAACGCAAAATCTTTAAGCATTAAGATTAAGATTAAGATTAAGATTAAGAT 52545  
 QY 4759 TAAACTAGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 4818  
 DB 52546 TAAACTAGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 52605  
 QY 4819 TACAAATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 4878  
 DB 52606 TACAAATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 52665  
 QY 4879 GTACATGTGATGTGTTGAGTGAATTTATTTATTTATTTATTTATTTATTTATTTAT 4938  
 DB 52666 GTACATGTGATGTGTTGAGTGAATTTATTTATTTATTTATTTATTTATTTATTTAT 52724  
 QY 4939 CGGCTCAGCAAGCTCTTAAAGGAGATGTAAGGAGATGTAAGGAGATGTAAGGAGAT 4988  
 DB 52725 CGGCTCAGCAAGCTCTTAAAGGAGATGTAAGGAGATGTAAGGAGATGTAAGGAGAT 52784  
 QY 4989 ----- 4988  
 DB 52785 CT 52844  
 QY 4989 -TAGCTGAGAGAGAGATGTAAGTGAATTAATTAATTAATTAATTAATTAATTAATTA 5047

DB 52845 CTCTGTTAAATGAAGATGATGTTCTAGTTAAATTAATTAATTAATTAATTAATTAAT 52904  
 QY 5048 CTATTTAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5107  
 DB 52905 CTATTTAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 52964  
 QY 5108 GGAAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5167  
 DB 52965 GGAAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53019  
 QY 5168 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5227  
 DB 53020 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53079  
 QY 5228 AATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5287  
 DB 53080 AATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53132  
 QY 5288 AGCTGTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5347  
 DB 53133 AGCTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 53192  
 QY 5348 GAAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5407  
 DB 53193 GAAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53243  
 QY 5408 AGCACTGAAAGCCGCTGAG 5467  
 DB 53244 AGCACTGAAAGCCGCTGAG 53302  
 QY 5468 ACAATGAGGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5527  
 DB 53303 ACAATGAGGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 53362  
 QY 5528 AGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5587  
 DB 53363 AGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53422  
 QY 5588 TGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5647  
 DB 53423 TGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53482  
 QY 5648 TGTGTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5707  
 DB 53483 TGTGTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 53542  
 QY 5708 AATTTTGAACCTTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5767  
 DB 53543 AATTTTGAACCTTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53602  
 QY 5768 TGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5827  
 DB 53603 TGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53662  
 QY 5828 ATGTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5887  
 DB 53663 ATGTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53720  
 QY 5888 TTGACCTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5947  
 DB 53721 TTGACCTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53780  
 QY 5948 TGAATTTTGA 5958  
 DB 53781 TGAATTTTGA 53791  
 RESULT 11  
 HSP63G01  
 LOCUS HSP63G01 835 bp DNA linear P01 04-JAN-2001  
 DEFINITION Homo sapiens P63 protein (P63) gene, exon 1.  
 ACCESSION AF124528







Query Match	Score	DB 2:	Length
Local Similarity	6.74;	402.2;	56268;
Matches 541; Conservative	68.34;	Prd. 1.6e-67;	
	0;	MemMatches 229;	Indels 22; Gaps 4;
*	28176	28880: contig of 705 bp in length	
*	28881	28980: gap of 100 bp	
*	28981	29685: contig of 705 bp in length	
*	29686	29785: gap of 100 bp	
*	29786	30479: contig of 694 bp in length	
*	30480	30579: gap of 100 bp	
*	30580	31254: contig of 675 bp in length	
*	31255	31354: gap of 100 bp	
*	31355	32042: contig of 688 bp in length	
*	32043	32142: gap of 100 bp	
*	32143	32847: contig of 705 bp in length	
*	32848	32947: gap of 100 bp	
*	32948	33553: contig of 706 bp in length	
*	33554	33753: gap of 100 bp	
*	33754	34469: contig of 716 bp in length	
*	34470	34569: gap of 100 bp	
*	34570	35281: contig of 712 bp in length	
*	35282	35381: gap of 100 bp	
*	35382	36076: contig of 695 bp in length	
*	36077	36176: gap of 100 bp	
*	36177	36887: contig of 711 bp in length	
*	36888	36987: gap of 100 bp	
*	36988	37684: contig of 697 bp in length	
*	37685	37784: gap of 100 bp	
*	37785	38484: contig of 700 bp in length	
*	38485	38584: gap of 100 bp	
*	38585	39274: contig of 690 bp in length	
*	39275	39374: gap of 100 bp	
*	39375	40069: contig of 695 bp in length	
*	40070	40169: gap of 100 bp	
*	40170	40880: contig of 711 bp in length	
*	40881	40980: gap of 100 bp	
*	40981	41780: contig of 720 bp in length	
*	41701	41800: gap of 100 bp	
*	41801	42513: contig of 713 bp in length	
*	42514	42613: gap of 100 bp	
*	42614	43309: contig of 656 bp in length	
*	43310	43409: gap of 100 bp	
*	43410	44112: contig of 703 bp in length	
*	44113	44212: gap of 100 bp	
*	44213	44914: contig of 702 bp in length	
*	44915	45014: gap of 100 bp	
*	45015	45722: contig of 708 bp in length	
*	45723	45822: gap of 100 bp	
*	45823	46541: contig of 719 bp in length	
*	46542	46641: gap of 100 bp	
*	46642	47355: contig of 714 bp in length	
*	47356	47455: gap of 100 bp	
*	47456	48171: contig of 716 bp in length	
*	48172	48271: gap of 100 bp	
*	48272	48954: contig of 683 bp in length	
*	48955	49054: gap of 100 bp	
*	49055	49774: contig of 720 bp in length	
*	49775	49874: gap of 100 bp	
*	49875	50604: contig of 730 bp in length	
*	50605	50704: gap of 100 bp	
*	50705	51409: contig of 705 bp in length	
*	51410	51509: gap of 100 bp	
*	51510	52214: contig of 705 bp in length	
*	52215	52314: gap of 100 bp	
*	52315	52991: contig of 677 bp in length	
*	52992	53091: gap of 100 bp	
*	53092	53795: contig of 704 bp in length	
*	53796	53895: gap of 100 bp	
*	53896	54628: contig of 733 bp in length	
*	54629	54728: gap of 100 bp	
*	54729	55450: contig of 722 bp in length	
*	55451	55550: gap of 100 bp	
*	55551	56268: contig of 718 bp in length	

Oy		4921	AGGTTAATTTCTTTCATTACTCGGGTCACGCAAAAGCTTCAGAAGGGAATGTGAAGGATAATCT	4980
Dd		27366	AAGCCANNN	2730
Oy		4981	CTTCTCCTTAGCTGAGAGAAAGTAGCTTCTAAGTTAAATATTAACAAGAAATTTCCCCT	5040
Dd		27306	NNA	2724
Oy		5041	GTCCTTGCTAATTTGAGAATTGTGCACCACMACGCCGTTGGCTGAAGAGGAAACTGAAGGG	5100
Dd		27246	GCTTTCCTCATTTAGATATGTGACCAACACAGCGCGTTGGCTGAAGAGGAGCAGAAAGGG	2718
Oy		5101	CGGGAGAGGAGCAAAATAGATGAABAAAAAACAAACCAAACAAATCTTCCCTAAGAGCTCT	5160
Dd		27186	GGTGGGGGTGGGG-----CHCGAGAACCCGGAABAAAABAABBAAGTCTCTAAGGTCCTC	2713
Oy		5161	ACCAAACTTTTATGCCCCCACAAATGTGCACAGAAATCTCAANTGMAACAGATATCCAGA	5220
Dd		27131	GCAAAACATTGTAGCAACAGACGAACATCAAGAGCTCTCAATCATCACTCGAANAACAGA	2707
Oy		5221	TACAAAGAAAGTTATATGATGTAGCTGAGCAGSAGSTGAAACATCATCACTCAAGTTCAGTA	5280
Dd		27071	TACAAAGAAATGTATATTCAGTTGAGAGCAAGGAGCAATTTATTAACCTCAGT-----GA	2701
Oy		5281	AAAGTCACGGGCTGCTGAATTAATCTGTATGTGCATTTAGCCAGATCCATCAACAGCA	5340
Dd		27018	CAAGTCTGGCTCTGTGATTTAACTGTATGCAATTCATTACAGACCAACCAATCCCAAGC	2695
Oy		5341	GAGATTAGAA GTTCAGAGATGCTCTCCAGCTCCAATTCGCAACAAAGTGTGGCTACTA	5400
Dd		26958	AAATATCAGAA GTTCAGAAATGCTCTCAAAATTCG-----CAACAAAGTGTGGCACTC	2690
Oy		5401	TAGCTCAAGACCTCTGAAAGCCGTGAGAGAGGAGGAGAAACAACAGTAGAGAGATGCCCA	5460
Dd		26907	TAGCTCAAGAGGCTCTAAACCTGTGGCAGA--GAGAGAGAAACA GCTTTACAGGSGGSCCA	2684
Oy		5461	GCTGTGAAGAAATCGAGTGTATTATGAGTTTTAGTCATTTAGATGAATATCATATGGCTAAA	5520
Dd		26848	CGTGTGAAGAAATTAAGCGGTTTATATGCTCTGGTACTTGAAAGACTCTCATATGGCTAAA	2678
Oy		5521	TCAAGAAAGCGTCCGCCCTCTTGCAAAATATGATGAAGAGAGAGAGTGCCTAAACCTCTA	5580
Dd		26788	GGAAGAAAGCGCCGCCCTCTTGCAAAATCTGAGTAAAGTAAAGGAGAGAGTGTCTAAACCTCTA	2672
Oy		5581	TGTCGTATGACATTGAACCTATATGCTTTAGCCTCCCGCTTATATATCTATATADAC	5640
Dd		26728	TGTCGTATGACATTGAACCTATATGCTTTAGCCTCCCTGCTATACATACCTAAATATTTCTC	2666
Oy		5641	AGGATTTTNGATATTTTATATATATATGTTCTCCGTCGTGATATCAAGA CAGTGTGA	5700
Dd		26668	AGGATTTATATGATATTTTATATGAATGCTCCCATCTGTGTGTATCAAGAAGAGTGA	2660
Oy		5701	GGAATGAATTT 5712	
Dd		26608	GGAATGAATTT 26597	
RESULT 14				
AF533892				
LOCUS	AF533892	208158 bp	DNA	linear ROD 31-JUL-2003
DEFINITION	Mus musculus p63 (Tip63) gene, alternatively spliced, complete cds.			
ACCESSION	AF533892			
VERSION	AF533892.1 GI:32812140			
SOURCE				
ORGANISM	Mus musculus (house mouse)			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Rodentia; Scurionomachi; Muridae; Murinae; Mus.				
REFERENCE				
AUTHORS	Ishikawa, J.E. and Zhang, M.Q.			
TITLE	Detection of polyadenylation signals in human DNA sequences			
JOURNAL	Gene 231 (1-2), 77-86 (1999)			

MEDLINE 99250252  
 PUBMED 10231571  
 REFERENCE 2 (bases 1 to 208158)  
 AUTHORS Kuf, J.S., Wang, J.H., Zhang, M.Q. and Mills, A.A.  
 JOURNAL Direct Submission  
 Submitted (02-AUG-2002) Mills Lab, Cold Spring Harbor Laboratory, 1  
 Bungtown Rd, Cold Spring Harbor, NY 11724, USA  
 FEATURES  
 source  
 1..208158  
 /organism="Mus musculus"  
 /mol\_type="Genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="16"  
 /map="19.30"  
 <1..208158  
 /gene="Trp63"  
 /note="Synonym: P63"  
 join<1..209,79122..79250,80020..80152,136373..136627,  
 178277..178463,179678..179793,181332..181441,  
 182160..182296,184252..184334,187083..187219,  
 198613..198770,200897..201041,202292..202385,  
 205091..208158)  
 /gene="Trp63"  
 /product="p63 TA alpha"  
 /note="alternatively spliced"  
 join<1..209,79122..79250,80020..80152,136373..136627,  
 178277..178463,179678..179793,181332..181441,  
 182160..182296,184252..184334,187083..187219,  
 198613..198770,200897..201041,205091..208158)  
 /gene="Trp63"  
 /product="p63 TA beta"  
 /note="alternatively spliced"  
 join<1..209,79122..79250,80020..80152,136373..136627,  
 178277..178463,179678..179793,181332..181441,  
 182160..182296,184252..184334,187083..187219,  
 192774..192828)  
 /gene="Trp63"  
 /product="p63 TA gamma"  
 /note="alternatively spliced"  
 <1..209  
 /gene="Trp63"  
 /note="transactivating isoform only"  
 /numbers1  
 <1..147  
 /gene="Trp63"  
 /note="transactivating isoform; partially characterized  
 using 5' RACE"  
 join1148..209,79122..79250,80020..80152,136373..136627,  
 178277..178463,179678..179793,181332..181441,  
 182160..182296,184252..184334,187083..187219,  
 198613..198770,200897..201041,202292..202385,  
 205091..205387)  
 /gene="Trp63"  
 /product="p63 TA alpha"  
 /codon\_start=1  
 /protein\_id="AA019682.1"  
 /db\_xref="GI:1281241"  
 /translation="MNEFTRSCATLTQVCPDPIQRFITPAHFWKSESYRSMOST  
 QTSFSEFSEFQHTWDELPDIPISVDPPIELNVDPSRGNKTKLEISMOCIKQSD  
 LADKMPQYTLGLNMQDQIQNGSSSTSPNTDQNSVTAPSPAGSFTDLAS  
 PSPAFSTSDITKGFHSFVDSFQDSSTAKATVSTETLKLYCQIATCPDIQKMP  
 PQGVALDAMPVYKKAHVEVYVAKRPNHLESRNEQSLAPSHLIRVGNSHAYV  
 LDRITRGSVLAVPEPQVGTETVTLVPMKNSACVGGNRRPILITVLETRDGV  
 EGRIFEGARICACPGARDKADSDISIRKQVSSAAGCGGTGRPRQNTICMTSICK  
 RSPSDILYLVPRGRETYEMALKIESLELQVLPQRTIETFRQOQOQOHLLKQ  
 TSMQSSQVYNSPPLKNNKSNKLPSVQLINPQQRNALPTTPMEGMANIPMGT  
 HNPAGDNGSLSPQALPPLSMPTSTQTPPTGCSIVSFLALGSGSLDFT  
 TQGLTITVQIHYEMDMLAKLPEQFRATVKGILDRHRLDLSPPHLITPSGAS  
 TVSVSSSTRGEVRIADVRLTQGITSPPRDNDMPDMSRANKQDRIKRRG"  
 join1148..209,79122..79250,80020..80152,136373..136627,  
 178277..178463,179678..179793,181332..181441,  
 182160..182296,184252..184334,187083..187219,  
 198613..198770,200897..201041,205091..205106)

/gene="Tnp63"  
 /codon\_start=1  
 /product="p63 TA beta"  
 /protein\_id="AAP87983.1"  
 /db\_xref="GI:32812142"  
 /translation="MNPETSRCAITLQYCPDPVYIQRPIETPAHFSKESYRNSAOSTRTQTEFLSEVPFQHIWDLPEQILCSVPDIELNPFDRSENGATNKEIISMDCIRNQBDLSIPMPMPQNTYGLNLSMDQIQNGSSSTSPVNTDHAQNSVAPSPAPQSPIDTALSPSAIPSPQNTYGLNLSMDQIQNGSSSTSPVNTDHAQNSVAPSPAPQSPIDTALSPPCAVIRAMPYKKAELVTEVVKCPNHELSEFNEGOIAPPSLIIVTEENSAQVYEDITGRGIVLPVPEPVQTEFTFTVYNMCSNVCUGMNRPLLIVTELETFDQVIGRCRCEALICACPRDKADEDSIKRQVDSAKNGDAFRNTHGIIOMTSIKRRASPDDELILVPRGRATEMLKIKSELMOYLPQHTIETVYRQOQQQHLLQKHLISA  
 CPENELVEVRGAEVTPQSDVFPFRHSNPNHSVPE"  
 79122. 79250  
 /gene="Tnp63"  
 /note="transactivating isoform only"  
 /number=2  
 80020. 80152  
 /gene="Tnp63"  
 /note="transactivating isoform only"  
 /number=3  
 join(117914. 118099,136373. 136627,178277. 178463, 179678. 179793,181332. 181441,182160. 182296, 184252. 184334,187083. 187219,198613. 198770, 200897. 201041,202292. 202385,205091. >208158)  
 /gene="Tnp63"  
 /product="p63 DN alpha"  
 /note="alternatively spliced"  
 join(117914. 118099,136373. 136627,178277. 178463, 179678. 179793,181332. 181441,182160. 182296, 184252. 184334,187083. 187219,198613. 198770, 200897. 201041,205091. 208158)  
 /gene="Tnp63"  
 /product="p63 DN beta"  
 /note="alternatively spliced"  
 join(117914. 118099,136373. 136627,178277. 178463, 179678. 179793,181332. 181441,182160. 182284, 184252. 184334,187083. 187219,192714. 192828)  
 /gene="Tnp63"  
 /product="p63 DN gamma"  
 /note="alternatively spliced"  
 117914. 118099  
 /gene="Tnp63"  
 /note="first exon in DN isoypes; spliced out of transactivating isoform"  
 /number=3  
 117914. 118057  
 /gene="Tnp63"  
 /note="dominant negative isoform, partially characterized using 5' RACE"  
 join(118058. 118099,136373. 136627,178277. 178463, 179678. 179793,181332. 181441,182160. 182296, 184252. 184334,187083. 187219,198613. 198770,

```

200897.. 201041,202292.. 202385,205091.. 205387)
/gene="Ttp63"
/codon_start=1
/product="p63 DN alpha"
/protein_id="AAP87986.1"
/db_xref="gi:32812144"
/translation="MLYLNNNAQTOFSEPOYTNLGLNSMDQIQNGSSSTSPYNTD
AONSVTAPSPQSPSTDAISBPALPSNTDVPSPHSFEDVSPQSSSTAKSATWTST
ELKLYCOIATKCPLOIKWTPPEQCAVIRAMVYKAAHYEVKRPENHLSREN
EGOLAPSHLIRVGNHQAQYVEDPITGROSVLVEPEPOVTEFTYLYKPFMNSSC
VGNMRPRLIIVLETRDQVIGRCEPARIACGRDRKADSDIRKQOVSAN
CDGTRPERONTGICOWTSIKRSPDDELIVYPRGRTYMLKIKESLMOYLP
QHTITRQOQOQHQLQKQTSNQSOSYNSSPKANKSMKLSQSLINPO
RNALPTTPEMGANIPMGTHPMAGNGLSPYQALPELSPSHCPRPPPIE
PNPMDSRNRKQRIKEGR"
join(118058..118099,136373..136627,178277..178463,
179678..179793,181332..181441,182160..182296,
184252..184334,187083..187219,198613..198770,
200897..201041,205091..205106)
/gene="Ttp63"
/codon_start=1
/product="p63 DN beta"
/protein_id="AAP87986.1"
/db_xref="gi:32812145"
/translation="MLYLNNNAQTOFSEPOYTNLGLNSMDQIQNGSSSTSPYNTD
AONSVTAPSPQSPSTDAISBPALPSNTDVPSPHSFEDVSPQSSSTAKSATWTST
ELKLYCOIATKCPLOIKWTPPEQCAVIRAMVYKAAHYEVKRPENHLSREN
EGOLAPSHLIRVGNHQAQYVEDPITGROSVLVEPEPOVTEFTYLYKPFMNSSC
VGNMRPRLIIVLETRDQVIGRCEPARIACGRDRKADSDIRKQOVSAN
CDGTRPERONTGICOWTSIKRSPDDELIVYPRGRTYMLKIKESLMOYLP
QHTITRQOQOQHQLQKQTSNQSOSYNSSPKANKSMKLSQSLINPO
RNALPTTPEMGANIPMGTHPMAGNGLSPYQALPELSPSHCPRPPPIE
TDCSIVRIMOV"
join(118058..118099,136373..136627,178277..178463,
179678..179793,181332..181441,182160..182296,
184252..184334,187083..187219,192714..192828)
/gene="Ttp63"
/codon_start=1
/product="p63 DN gamma"

Query Match 46: Score 273.4; DB 10; Length 208158;
Best Local Similarity 81.8%; Freq. NO. 2.7e-42;
Matches 328; Conservative 0; Mismatches 71; Indels 2; Gaps 1;

QY 5558 GGAGAGAGTGCCTTAACCTTCTATGCTGATGATGACCTTATGCTTTCCTCC 5617
DB 1 GGAGGAGAGTGTCTAACTTCTATGCTGATGATGACCTTATGCTTTCCTCC 60
QY 5618 CCGCTTATATCTATATATACAGAGTATTTGTGATATTTTATATATATGTTCTCCGT 5677
DB 61 TGCGTACATACCTAGATATCTCAGGTATATGATATTTTATATAGATTTGCTCCATC 120
QY 5678 CGTGTATATCAAGACAGTTGAGGAATGAAATTTGAAACTTCAGGTGTCACCTTA 5737
DB 121 TGTGTGATCAAGAGAGTTGAGGAATGAAATTTGAAACTTCAGGTGTCACCTTA 180
QY 5738 CAGTACTGACCTGACCTTACATCCAGCGGTGAGTTGAAATGTCACATTACTTCTCA 5797
DB 181 CAGTACTGACCTGACCTTACATCCAGCGGTGAGTTGAAATGTCACCTGACCTTCA 240
QY 5798 AACTTATGAGTGTGCTGTGATATGATGATGATGATGATGATGATGATGATGATG 5857
DB 241 AACTGAACTGTGTGCTGTGATATGAGTGTGTGCTGTGTAC--TGAAGAAAGCC 298
QY 5858 TCTTGTGTGATGACGACAGTATATTTTGTGACCTTCTGAGATTTTAAATGCTG 5917
DB 299 TCTTGTGTGACGACAGTATATTTTGTGACCTTCTGAGATTTTAAATGCTG 358
QY 5918 TGCACATATTTTGTGATGCTTTTGTGATATTTGGA 5958
DB 359 TAGATATCTTTTGTGACGATATCTTATATTTCTTGA 399

```

## RESULT 15

AC103882/c

LOCUS

DEFINITION

AC103882

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

AC103882 159249 bp DNA linear pri 15-OCT-2002  
 Homo sapiens BAC clone RP11-733G6 from 2, complete sequence.  
 AC103882  
 AC103882.5 GI:23130727  
 HTG.  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Sulston, J.E. and Waterston, R.  
 Toward a complete human genome sequence  
 Genome Res. 8 (11), 1097-1108 (1998)  
 9847074  
 2 (bases 1 to 159249)  
 Wang, C. and Bielik, L.  
 The sequence of Homo sapiens BAC clone RP11-733G6  
 Unpublished (2001)  
 3 (bases 1 to 159249)  
 Waterston, R.H.  
 Direct Submission  
 Submitted (29-NOV-2001) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 4 (bases 1 to 159249)  
 Waterston, R.H.  
 Direct Submission  
 Submitted (30-MAR-2002) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 5 (bases 1 to 159249)  
 Waterston, R.H.  
 Direct Submission  
 Submitted (18-SEP-2002) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 6 (bases 1 to 159249)  
 Waterston, R.H.  
 Direct Submission  
 Submitted (03-OCT-2002) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 7 (bases 1 to 159249)  
 Waterston, R.H.  
 Direct Submission  
 Submitted (15-OCT-2002) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Sep 16, 2002 this sequence version replaced gi:19703341.  
 ----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc  
 Contact: sapiens@wustl.edu  
 ----- Summary Statistics  
 Center Project name: H.NH073306  
 -----  
 NOTICE: This sequence may not represent the entire insert of this  
 clone sections once, or longer because we provide a small overlap  
 between neighboring data submissions.  
 This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate  
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by sequence  
 from more than one subclone; and the assembly was confirmed by  
 restriction digest.

## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McChesron, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

## SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Prengen, E., Tatem, M., Catane, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.choir.org>

VECTOR: pBAC3.6

## NEIGHBORING SEQUENCE INFORMATION:

This sequence is not the entire repeat insert of the clone. This clone is overlapped by AC11754 and AC09656.

There is a simple sequence repeat from base 4648 to 46739. This region does not meet required finishing standards.

## FEATURES

## SOURCE

Location/Qualifiers  
1..159249  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/map="2"  
/clone="RP11-733G6"  
/clone\_lib="RPCI-11"  
102..178  
/rpt\_family="LI"  
178..571  
/rpt\_family="MALR"  
572..626  
/rpt\_family="LI"  
765..942  
/rpt\_family="LI"  
1024..1916  
/rpt\_family="LI"  
1917..2217  
/rpt\_family="ALU"  
2218..2499  
/rpt\_family="LI"  
2499..33481  
/rpt\_family="LI"  
3482..3782  
/rpt\_family="ALU"  
3783..3929  
/rpt\_family="LI"  
3930..4251  
/rpt\_family="ALU"  
4252..4764  
/rpt\_family="LI"  
4765..4986  
/rpt\_family="LI"  
4994..5023  
/rpt\_family="(TA)n"  
5024..5176  
/rpt\_family="LI"  
5236..5404  
/rpt\_family="LI"  
5520..5705  
/rpt\_family="LI"  
5747..5777  
/rpt\_family="AT\_rich"  
5782..5972  
/rpt\_family="LI"  
6450..6474  
/rpt\_family="(TG)n"  
6998..7244  
repeat\_region

/rpt\_family="MIR"  
7948..8299  
/rpt\_family="MER2\_type"  
8868..9128  
/rpt\_family="MER2\_type"  
9299..9510  
/rpt\_family="ERV1"  
9516..9711  
/rpt\_family="ERV1"  
9717..10011  
/rpt\_family="ALU"  
10275..10313  
/rpt\_family="AT\_rich"  
10375..10528  
/rpt\_family="(TA)n"  
10529..10659  
/rpt\_family="(TATAA)n"  
10695..11412  
/rpt\_family="ERV1"  
11440..11544  
/rpt\_family="ERV1"  
11569..11927  
/rpt\_family="MALR"  
12292..12495  
/rpt\_family="ERV1"  
12856..13041  
/rpt\_family="ERV1"  
13197..14809  
/rpt\_family="LI"  
14829..14876  
/rpt\_family="MIR"  
14898..15187  
/rpt\_family="MER1\_type"  
15243..15523  
/rpt\_family="ALU"  
15728..15999  
/rpt\_family="ALU"  
16186..16418  
/rpt\_family="MALR"  
16727..17022  
/rpt\_family="ALU"  
17165..17468  
/rpt\_family="ALU"  
17670..17868  
/rpt\_family="MIR"  
18530..18674  
/rpt\_family="MIR"  
19048..19069  
/rpt\_family="AT\_rich"  
20886..21207  
/rpt\_family="CR1"  
21379..22209  
/rpt\_family="LI"  
22342..22558  
/rpt\_family="MIR"  
22572..22872  
/rpt\_family="CR1"  
23219..23264  
repeat\_region

Query Match 3.6%; Score 212.4; DB 9; Length 159249;

Best Local Similarity 63.3%; Pred. No. 1.2e-30;

Matches 413; Conservative 0; Mismatches 221; Indels 18; Gaps 5;

QY 3293 TGCCTGCTCATGCTGCTATATCCCAACACTTGGCAAGCCGCTCGGAGAGATAGCTTGA 3352  
DB 85766 TGGTGGCTTCTGCTTAACCTCCAGACACTTGGAGGCAAGTGGAGGATCTTGA 85707  
QY 3353 TCACGAGTTGAGACCACTGAGGCAACACAGCAAGACCCATATCTAATAAACAATA 3412  
DB 85706 CTCAGGCAATTGAGACCACTGAGGCAATATAAAGGCTCTGCTATCTAATAATTA 85647  
QY 3413 CAAACAAATTAACCTGGGTATGCTGCTACCTGATGCAAGCTACAGCAAGGCT 3472

Db 85646 -----AAATTACTGGGTGTGTGTACATGCTGTACTCCAGTACTCAGAGGCT 85593  
Qy 3473 GAGCGAAGAGATCACTTGAAGCCGAGAGGTTGAGGCTGCACTGATCCATGAACCGCTG 3532  
Db 85592 GAAGGGAGAGATCACTGAGCTCGAGCCGAGAGGCTGAGGCTGAGCCAGATGGCAACA 85533  
Qy 3533 CTACACT-CACTGTGGTGAAGTGCAGAGAGCTGTCTCNAATAATTAATTAATAA 3591  
Db 85532 CTGTACTCAGCTGGGTGACAGACAGCAAGCTCTGTCTCAAAAAAAAAAAAAAG 85473  
Qy 3592 ATACTTTAAAAAAGAAAAATTAATTAATTAATAACACACACTAGAT--GT 3649  
Db 85472 AAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGATGTTAA 85413  
Qy 3650 TTGCAATTGATTATTGGAGTCTATATCCCTGAAAGTTAAATTAATAATTAGAGA 3709  
Db 85412 AAGCAATTGTTCCTCGAATACTATATCACTGTTCTGTATCACTAGCAGA 85353  
Qy 3710 GTTCTTCCTCATTTCTGAGAGCTGCAATTGTAAATATCAGAGCTAGAGGAACACTAG 3769  
Db 85352 AAACAAGCCTTCACTTACGAAACTGGGTTGAGGAGAGCAAGCAAGCAAGCAAGAA 85293  
Qy 3770 GGCTGGCACTCCAAAGTGTGTCCAAAGACACAGACATCAAGTAACCTGGAAAGTGT 3829  
Db 85292 GGTGCTAGGAAAGAACTGAGAAATGTGTGACAGCAC-----TGCTGGAGCTTGT 85240  
Qy 3830 TAGAATGAGAGTCTTGAAGCTTACCCAGACCTACTGAACA--GAATCTGCATTAC 3887  
Db 85239 CAGAAATGCAAGATCTTGGGCTTCTCCCAAGACCACTGAATCAGGGCATGTATTTAAC 85180  
Qy 3888 AAGATTTCTAGTGGCTCAGCGGACATTAATACTTGAAGAGCTTGCACTA 3939  
Db 85179 AAGACCTCAGGTGATCTCTGACAGGTTAAACTTGAAGACCTGACAA 85128

Search completed: September 27, 2004, 14:32:31  
Job time : 22298 secs





Patent No. 6531423  
GENERAL INFORMATION: Karl et al  
APPLICANT: GUEGLER, ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CL000636  
CURRENT APPLICATION NUMBER: US/09/738,894A  
CURRENT FILING DATE: 2000-12-18  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 36651  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(36651)  
OTHER INFORMATION: n = A,T,C or G  
US-09-738-894A-3

Query Match	3.1%;	Score 182.2;	DB 4;	Length 36651;
Best Local Similarity	76.6%;	Pred. No. 1.1e-31;		
Matches 249;	Conservative 0;	Mismatches 73;	Indels 3;	Gaps 2;

Qy	3299	GTGCTGTGGCTCATGCTCTGTAATCCCAACAATTTGGGAAACCGGGTGGGAGGTGACCT	3349
Db	11525	GTGCTGTACCTACGCTCTGTAATCCCAACCTTTGGGAAACCGAAGTGGGCAATCCCTT	11584
Qy	3350	GAGTCACACATTTTGAGACACAGTCACGGGACACAGCAGAGACCCCATATCAAAAAAAC	3409
Db	11585	GAGGCGACAAAGTTTGAGACCGGTCTAGGCAACATAGCAAGACTTTGTCTTAAAAA--A	11642
Qy	3410	AAACAAACCAAAATTTACTCTGGATGTGTGTGTCTACCTGTATGTCACAGTACACAGGAA	3469
Db	11643	AAAAAAAAAAAAATTTACTCTGGCATGTGTGTATGACACTGTATATCCAGCTACTCAAGAG	11702
Qy	3470	GCTAGAGCAAGAAAGATCACTTAAAGCCACAGAGTTTAAAGCTTCAGTGATCATGAACCG	3529
Db	11703	GCTAGAGCAAGAGATGCTTTAGACCTCTGGAGATTGAAGCTGCAAGGTGACCGTGATATCA	11762
Qy	3530	CTGCTACACT--CAGTCTGGGTGAACAGTCAGAAAGCTGTCTCAAAAAATTAATTAATAATA	3588
Db	11763	CCACGTGCACTCACGCTCGATGTGACAGGCGAAGACCTGTCTCTAAACAAACAAACAAAC	11822
Qy	3589	AAAAACACTTTAAAAAACAATAAT	3613
Db	11823	AAAAACGAAACAAAAACAATATAT	11847

```

RESULT 3
US-09-964-469-3
Sequence 3, Application US/09964469
Patent No. 6579709
GENERAL INFORMATION:
APPLICANT: GUSELIER, Karl et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THREEOF
FILE REFERENCE: CLO00636DIY
CURRENT APPLICATION NUMBER: US/09/964,469
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: 60/209,331
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: 09/738,894
PRIOR FILING DATE: 2000-12-18
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 36651
TYPE: DNA
ORGANISM: Human
FEATURE:

```

```

; NAME/KEY: misc_feature
; LOCATION: (1) .. (36651)
; OTHER INFORMATION: n = A,T,C or G
US-09-964-469-3

```

Query Match	3.14	Score 182.2	DB 4	Length 36651
Best Local Similarity	76.64	Pred. No. 1.1e-31		
Matches 249	Conservative 0	Mismatches 73	Indels 3	Gaps 2

Oy	3290	GGTGGTGGGCTCAGGCGCTGTAATCCCAACCTTTGGGAAAGCGGGATCCGGAGAGTAACTT	3349
Db	11525	GTGCTGTACTCAGCGCTGTATCCCAACCTTTGGGAAAGCCAAAGTGGAGAAATCCCTT	11584
Oy	3350	GAGTCACAGCATGTTGAGACCGACAGTCCAGGCGCAACAGACAGACCCCATTCCTAAAAAACA	3409
Db	11585	GAGGCGACAGATGTTGAGACCGCTCTAGGCGACATGACCAAGACTTTGCTCTAAAAA-A	11644
Oy	3410	AAACAAAACAAAATTAACCTGGGATATGGTGTCTCACTCGTATGTCGAAAGCTACACAGGAA	3469
Db	11643	AAAAAAAAAAAAAAAAATTACTCTGGCGATATGATATGCACTCTGTAATCCAGCTACTCAGAG	11702
Oy	3470	GGTGGGCGAAGGATCACTTGAAGCCCAAGAGGTTGAGGCTCGAGTATCAAGAAAGCTG	3529
Db	11703	GCTGAAGCGAGAGATTCCTTGAACCTCGAGAGTTTGATGGCTGACAGTACGCGATATCA	11762
Oy	3530	CTGCTTAACCTCACTTCGCTCGGTCAACAGTCGCAAGAGCTGTCTCAAAATTAATTAATTAATTA	3588
Db	11763	CACATGCACTCAGCGCTGATGAGACAGGACAGCACTGTCTCTAAACAAACAAACAAACA	11822
Oy	3589	AAATTAACCTTTAAAAAACAATAAT	3613
Db	11823	AAAAAAAAACAAAACAAATATC	11847

RESULT 4  
US-09-301-665-3

```

? Patent No 6207876
? GENERAL INFORMATION:
? APPLICANT: KELLEMS, RODNEY E.
? APPLICANT: DAYTA, SURJIT K.
? APPLICANT: BLACKBURN, MICHAEL R.
? TITLE OF INVENTION: ADEONISINE DEAMINASE DEFICIENT TRANSGENIC MICE AND
? FILE REFERENCE: UTS# 243
? CURRENT APPLICATION NUMBER: 06/09/301, 665
? EARLIER FILING DATE: 1999-04-28
? EARLIER APPLICATION NUMBER: 60/083, 408
? EARLIER FILING DATE: 1998-04-29
? EARLIER APPLICATION NUMBER: 60/083, 370
? NUMBER OF SEQ ID NOS: 4
? SOFTWARE: Patentin Ver. 2.0
? SEQ ID NO 3
? LENGTH: 36741
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-301-665-3

```

Query Match	3.04	Score 176;	DB 3;	Length 36741;
Best Local Similarity	71.84	Pred. No. 2.9e-30;		
Matches 244; Conservative	0;	Mismatches 95;	Indels 1;	Gaps 1;

Oy 3257 TAAACCTTTAAATTTTAAAAAATAGATNGATGTGGGGCTCATGCCCTGATATCCA 3316  
 Db 4879 TTACCTCTTAAGAAATAGATGTGTATGTGGCCACGAGTGTGCTCAGCTCTGATATCCA 4938  
 Oy 3317 ACACTTGTGGAGACCGGATCGAGAGCATTACTTCCAGATTTGATGAGACCATGATCAG 3376  
 Db 4939 GCACCTTGTAGCCTTAGCGCGGAGATCATTTGATCCAGATTTGATGAGCATGAGCTGG 4998  
 Oy 3377 GCACACGACATACCCATATCTAAAAAACAACAAAAATTAACCTGGATAG 3435

Thu Sep 30 11:51:28 2004

us-10-030-294-2.std.inl

**Page 3**

[illegible]

```

RESULT 5
US-09-798-096-10/c
Sequence 10. Application US/09798096
Patent No. 6399378
GENERAL INFORMATION:
APPLICANT: Donna T. Ward
APPLICANT: Andrew T. Walt
TITLE OF INVENTION: AMTISENSE MODULATION OF RECQ2 EXPRESSION
FILE REFERENCE: RTS-0207
CURRENT APPLICATION NUMBER: US/09/798,096
CURRENT FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 89
SEQ ID NO 10
LENGTH: 99500
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-09-798-096-10

```

	Query Match	Best Local Similarity	7.2 %;	Score 115.4 ;	DB 4 ;	Length 99500 ;
	Matches 255 ;	Conservative 0 ;	Mismatches 86 ;	Indels 71 ;	Gaps 2 ;	
OY	3276	AAAAAAAAATGATGAAGTGTGGTGCTGCATCGCTGTAAATCCCAACACTTTGGGAAGCCCGGT	3333			
Dd	8073	AAAAAATTGATGAGAATGAAGATGGTCATGCTTTGAATCCAGCATTTGGAGGCCAMG	8014			
OY	3336	CGGAGAGATAGCTTAGTGAGTCCAGCAGTTTGAGACCAGTCCAGGCAACAACAGAACCCCA	3395			
Dd	8013	CAGAGATAGCTTAGTGAGCCAGAGATTGGAACAGGCTTAGGCAACAAGCAAAGACCCAG	7955			
OY	3396	TATCTATAAAAACAAACCAAAAACAATTAACCTGGGTATGGTGTGCTACCTGTATGTC	3455			
Dd	7953	TATCTACAAAATAA-----AAGGTTAGCCAAGATGGTGTACCTCTGTATGCT	7900			
OY	3456	AAGCTACACAGGAAGCTTAGGCAAGAGATCACTTGAGCCAGGAGGTTAGGCTGCACT	3515			
Dd	7899	CACCTTCTCGGAGGCTGAGGCAAGATCATCTGAGCCAGAGGTTAGAGGCTGCACT	7840			
OY	3516	GATCCATGAAGCGGCTGCTACACT-CAGCTGGGGGAGCAATGCAAGAASTGTCTCAAA	3574			
Dd	7839	GAGCAAGATCATTTACTAGCACTCCAGCTGTGATAGCTAGCAAGACCTGTCTGAABA	7788			
OY	3575	ATATATTAATATAAATAATPACTTTTAAAAACAATAATPATTA	3621			
Dd	7779	ATATATAAATACATCAATTAAGTAAATATATAAATAAATAAATAA	7733			

RESULT 6  
US-09-791-211-10/C  
; Sequence 10, Application US/09791211  
; Patent No. 6448080  
; GENERAL INFORMATION:  
; APPLICANT: Donna T. Ward  
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION  
; FILE REFERENCE: RTS-0205  
; CURRENT APPLICATION NUMBER: US/09/791.211

```

1 CURRENT FILING DATE: 2001-02-23
2 NUMBER OF SEQ ID NOS: 90
3 SEQ ID NO. 10
4 LENGTH: 98844
5 TYPE: DNA
6 ORGANISM: Homo sapiens
7 FEATURE:
8 NAME/KEY: unsure
9 LOCATION: 24962
10 OTHER INFORMATION: unknown
11 NAME/KEY: unsure
12 LOCATION: 64383
13 OTHER INFORMATION: unknown
14 NAME/KEY: unsure
15 LOCATION: 65468
16 OTHER INFORMATION: unknown
17 NAME/KEY: unsure
18 LOCATION: 65469
19 OTHER INFORMATION: unknown
20 NAME/KEY: unsure
21 LOCATION: 65470
22 OTHER INFORMATION: unknown
23 NAME/KEY: unsure
24 LOCATION: 65471
25 OTHER INFORMATION: unknown
26 NAME/KEY: unsure
27 LOCATION: 87130
28 OTHER INFORMATION: unknown
29 NAME/KEY: unsure
30 LOCATION: 89049
31 OTHER INFORMATION: unknown
32 OTHER INFORMATION: unknown
33 US-09-791-211-10

```

	Qy	3276	AAAAAAAAATGATGATGTGGTCTCATGTCAGTCCCTTAATTCCCAACACTTTGGGAAGCGGGT	3335
	Dd	44540	AAAGATATGGCCAGGATATGGTGGTTAACCCCTGTAAATCCAGACTTTGAAAAGTCAMG	4448R
	Qy	3336	CGGAGGATATGCTTGAGTCCAGCAGTTTGGACACCATGCAGGAGCACAGCAAGACC	3399S
	Dd	44480	CAGGAGGACCTGCTTAGACTAGAGGTTTAAAGACCAACCTGTGACACAG-----CA	4442Z
	Qy	3396	TATCTAAAAAACAACAACAATACTGGGATATGGTGTCTCACTGATAGTC	3455S
	Dd	44428	TGCTCTCAAAAAAABAAAAATTTAAAAATTAGCTGGCGTTGGTGGCAACAACCTGTATGCC	4430S
	Qy	3456	AAGCTTACAGGACCTGACGACAGAGATCATTTGAGCCGAGGAGTTGAGCTGCACT	3515S
	Dd	44368	GAGCTAATCTACAGAGACTGACGACAGAGATCATCTTAGTGCAGAGATTTGGAGTTAGCGT	4430S
	Qy	3516	GATCCATGAAACGGCTGTCTACAC--TCAGTCTGGGCTGACAGTGCANAAGACTGTCTCAAAA	3574S
	Dd	44308	GAGCTATGATTCGCATCACTGCACCTTGACCGCGGTAAAGTTCAAGAACCCTGTCTCTAAG	4422Z
	Qy	3575	ATTAATTAATTAATTAATTAATCTTTAAAAAAGAAAAATTAAT	3617
	Dd	44248	AAAAACAAAAATAAAAATTAATTTAAAAATCAAATAAT	44206

```

RESULT 7
US-09-671-317-485
; Sequence 485, Application US/09671317
; Patent No. 6528260
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bouguetelret, Lydie
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM

```

```
FILE REFERENCE: 62, US3, CIP
CURRENT APPLICATION NUMBER: US/09/671,317
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US 09/536,178
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/IB00/00403
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 60/126,269
PRIOR FILING DATE: 1999-03-25
PRIOR APPLICATION NUMBER: US 60/131,961
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 977
SOFTWARE: Patent.pm
SEQ ID NO 485
LENGTH: 49312
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5466..7466
OTHER INFORMATION: 5' regulatory region
NAME/KEY: exon
LOCATION: 7467..7725
OTHER INFORMATION: exon 1
NAME/KEY: exon
LOCATION: 20256..20355
OTHER INFORMATION: exon 2
NAME/KEY: exon
LOCATION: 36905..36975
OTHER INFORMATION: exon 3
NAME/KEY: exon
LOCATION: 45167..45248
OTHER INFORMATION: exon 4
NAME/KEY: exon
LOCATION: 45728..45965
OTHER INFORMATION: exon 5
NAME/KEY: misc_feature
LOCATION: 45966..49312
OTHER INFORMATION: 3' regulatory region
NAME/KEY: allele
LOCATION: 7564
OTHER INFORMATION: 10-286-289 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 7619
OTHER INFORMATION: 10-286-345 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 7649
OTHER INFORMATION: 10-286-375 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 17256
OTHER INFORMATION: 12-425-57 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 21590
OTHER INFORMATION: 12-421-135 : insertion of T
NAME/KEY: allele
LOCATION: 21595
OTHER INFORMATION: 12-421-140 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 36971
OTHER INFORMATION: 10-523-232 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 45214
OTHER INFORMATION: 10-289-201 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 45741
OTHER INFORMATION: 10-290-37 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 46029
OTHER INFORMATION: 10-290-326 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 46032
OTHER INFORMATION: 10-290-328 : deletion of G
NAME/KEY: primer_bind
LOCATION: 7276..7294
OTHER INFORMATION: 10-286.pu
NAME/KEY: primer_bind
LOCATION: 7676..7694
OTHER INFORMATION: 10-286.rp complement
NAME/KEY: primer_bind
LOCATION: 16839..16856
OTHER INFORMATION: 12-425.rp
NAME/KEY: primer_bind
LOCATION: 17297..17314
OTHER INFORMATION: 12-425.pu complement
NAME/KEY: primer_bind
LOCATION: 21455..21474
OTHER INFORMATION: 12-421.pu
NAME/KEY: primer_bind
LOCATION: 21886..21906
OTHER INFORMATION: 12-421.rp complement
NAME/KEY: primer_bind
LOCATION: 36740..36758
OTHER INFORMATION: 10-523.pu
NAME/KEY: primer_bind
LOCATION: 36997..37015
OTHER INFORMATION: 10-523.rp complement
NAME/KEY: primer_bind
LOCATION: 45020..45037
OTHER INFORMATION: 10-289.pu
NAME/KEY: primer_bind
LOCATION: 45413..45432
OTHER INFORMATION: 10-289.rp complement
NAME/KEY: primer_bind
LOCATION: 45705..45724
OTHER INFORMATION: 10-290.pu
NAME/KEY: primer_bind
LOCATION: 46104..46123
OTHER INFORMATION: 10-290.rp complement
NAME/KEY: primer_bind
LOCATION: 7545..7563
OTHER INFORMATION: 10-286-289.mis
NAME/KEY: primer_bind
LOCATION: 7565..7583
OTHER INFORMATION: 10-286-289.mis complement
NAME/KEY: primer_bind
LOCATION: 7600..7618
OTHER INFORMATION: 10-286-345.mis
NAME/KEY: primer_bind
LOCATION: 7620..7638
OTHER INFORMATION: 10-286-345.mis complement
NAME/KEY: primer_bind
LOCATION: 7630..7648
OTHER INFORMATION: 10-286-375.mis
NAME/KEY: primer_bind
LOCATION: 7650..7668
OTHER INFORMATION: 10-286-375.mis complement
NAME/KEY: primer_bind
LOCATION: 17239..17257
OTHER INFORMATION: 12-425-57.mis
NAME/KEY: primer_bind
LOCATION: 17259..17277
OTHER INFORMATION: 12-425-57.mis complement
NAME/KEY: primer_bind
LOCATION: 21576..21594
OTHER INFORMATION: 12-421-140.mis
NAME/KEY: primer_bind
LOCATION: 21596..21614
OTHER INFORMATION: 12-421-140.mis complement
NAME/KEY: primer_bind
LOCATION: 36952..36970
OTHER INFORMATION: 10-523-232.mis
NAME/KEY: primer_bind
LOCATION: 36972..36990
OTHER INFORMATION: 10-523-232.mis complement
NAME/KEY: primer_bind
LOCATION: 45195..45213
```



APPLICANT: DiFrancesco, Valentina  
APPLICANT: Beasley, Ellen  
TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
FILE REFERENCE: C001194  
CURRENT APPLICATION NUMBER: US 09/820,002  
CURRENT FILING DATE: 2001-03-29  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 21784  
TYPE: DNA  
ORGANISM: HUMAN  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(21784)  
OTHER INFORMATION: n = A,T,C or G  
US-09-820-002-3

Query Match 2.9%; Score 171.2; DB 4; Length 21784;  
Best Local Similarity 71.8%; Freq. No. 2,9e-25;  
Matches 254; Conservative 0; Mismatches 93; Indels 7; Gaps 2;

QY 3262 CTTTAAATTTTAAATAATAGATGAGTGTGTGCTCAATGCTTAATCCCAACT 3321  
DB 4238 CTACTAAAAATTTAAATAATAGAGTGTGTGTGCTCAATGCTTAATCCCAACT 4179  
QY 3332 TTGGAGACCGGCGGAGAGTAGCTTGAAGCCAGAGTTGAGACAGTCAAGGCGAC 3381  
DB 4178 TTGGAGACCGGCGGAGAGTAGCTTGAAGCCAGAGTTGAGACAGTCAAGGCGAC 4119  
QY 3382 ACAGCAAGACCCCATATCTTAAATAATCAAAACAAATAATCTGCTGATGCTGTG 3441  
DB 4118 ATGTGTAACCCCGCTCTA-----CTAAATAATCAAAATAATGCTGCTGTGCGCA 4065  
QY 3442 CTCACCTGTAGTCCAGACTACAGAGAGCTGAGGAGAGTCACTTGAAGCCAGAG 3501  
DB 4064 CGTGCCTGTAGTCTGAGCTACTAGAGAGCTGAGGAGAGATCACTTGAAGCTGAG 4005  
QY 3502 GTTAGAGCTGAGTGTGCTGAGAGAGCTGCTGACT-CACTGCTGCTGAGAGTCAAG 3560  
DB 4004 ATGAGAGTGTGAGTGTGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 3945  
QY 3561 AAGCTGTCTCAAAATATATATATATATATATATATATATATATATATATAT 3614  
DB 3944 ACTGTGCTCAAAATATATATATATATATATATATATATATATATATATAT 3891

RESULT 10  
US-09-750-580-1/c  
Sequence 1, Application US/09750580  
Patent No. 6455280  
GENERAL INFORMATION:  
APPLICANT: Yen, Frances  
APPLICANT: Demison, Blake  
APPLICANT: Bour, Barbara  
APPLICANT: Binau, Bernard  
APPLICANT: Dumas Milne Edwards, Jean-Baptiste  
APPLICANT: Duclet, Aymeric  
APPLICANT: Bouquelere, Lydie  
APPLICANT: Ebbels-Reed, Dana  
APPLICANT: Salter-Cid, Luisa  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH  
FILE REFERENCE: 89 US2 CIP  
CURRENT APPLICATION NUMBER: US 09/750,580  
CURRENT FILING DATE: 2000-06-21  
PRIOR APPLICATION NUMBER: US 09/599,362  
PRIOR FILING DATE: 2000-06-21  
PRIOR APPLICATION NUMBER: PCT/IB00/0101  
PRIOR FILING DATE: 2000-06-21  
PRIOR APPLICATION NUMBER: PCT/IB99/02058  
PRIOR FILING DATE: 1999-12-20

PRIOR APPLICATION NUMBER: US 49/469/099  
PRIOR FILING DATE: 1999-12-21  
PRIOR APPLICATION NUMBER: US 60/113,686  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: US 60/141,032  
PRIOR FILING DATE: 1999-06-25  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: Patent.pm  
SEQ ID NO 1  
LENGTH: 81001  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 10946..12946  
OTHER INFORMATION: 5'regulatory region  
NAME/KEY: exon  
LOCATION: 12947..12958  
OTHER INFORMATION: exon 1  
NAME/KEY: exon  
LOCATION: 13470..13526  
OTHER INFORMATION: exon 2  
NAME/KEY: exon  
LOCATION: 13641..13752  
OTHER INFORMATION: exon 3  
NAME/KEY: exon  
LOCATION: 14271..15968  
OTHER INFORMATION: exon 4  
NAME/KEY: misc.feature  
LOCATION: 15969..17969  
OTHER INFORMATION: 3'regulatory region  
NAME/KEY: allele  
LOCATION: 1239  
OTHER INFORMATION: 20-828-311 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 12347  
OTHER INFORMATION: 17-42-319 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 15241  
OTHER INFORMATION: 17-41-250 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 42218  
OTHER INFORMATION: 20-841-149 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 45442  
OTHER INFORMATION: 20-842-115 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 77058  
OTHER INFORMATION: 20-853-415 : polymorphic base C or T  
NAME/KEY: primer\_bind  
LOCATION: 929..949  
OTHER INFORMATION: 20-828-pu  
NAME/KEY: primer\_bind  
LOCATION: 1157..1177  
OTHER INFORMATION: 20-828-tp complement  
NAME/KEY: primer\_bind  
LOCATION: 12029..12050  
OTHER INFORMATION: 17-42-pu  
NAME/KEY: primer\_bind  
LOCATION: 12581..12603  
OTHER INFORMATION: 17-42-tp complement  
NAME/KEY: primer\_bind  
LOCATION: 14997..15012  
OTHER INFORMATION: 17-41-pu  
NAME/KEY: primer\_bind  
LOCATION: 15460..15482  
OTHER INFORMATION: 17-41-tp complement  
NAME/KEY: primer\_bind  
LOCATION: 42070..42090  
OTHER INFORMATION: 20-841-pu  
NAME/KEY: primer\_bind  
LOCATION: 42572..42591  
OTHER INFORMATION: 20-841-tp complement

NAME/KEY: primer bind  
 LOCATION: 45328..45347  
 OTHER INFORMATION: 20-842.pu  
 NAME/KEY: primer\_bind  
 LOCATION: 45863..45883  
 OTHER INFORMATION: 20-842.fp complement  
 NAME/KEY: primer\_bind  
 LOCATION: 76644..76664  
 OTHER INFORMATION: 20-853.pu  
 NAME/KEY: primer\_bind  
 LOCATION: 77166..77185  
 OTHER INFORMATION: 20-853.fp complement  
 NAME/KEY: primer\_bind  
 LOCATION: 1220..1238  
 OTHER INFORMATION: 20-828-311.mis  
 NAME/KEY: primer\_bind  
 LOCATION: 1240..1258  
 OTHER INFORMATION: 20-828-311.mis complement  
 NAME/KEY: primer\_bind  
 LOCATION: 12328..12346  
 OTHER INFORMATION: 17-42-319.mis  
 NAME/KEY: primer\_bind  
 LOCATION: 12348..12366  
 OTHER INFORMATION: 17-42-319.mis complement  
 NAME/KEY: primer\_bind  
 LOCATION: 15222..15240  
 OTHER INFORMATION: 17-41-250.mis  
 NAME/KEY: primer\_bind  
 LOCATION: 15242..15260  
 OTHER INFORMATION: 17-41-250.mis complement  
 NAME/KEY: primer\_bind  
 LOCATION: 42199..42217  
 OTHER INFORMATION: 20-841-149.mis  
 NAME/KEY: primer\_bind  
 LOCATION: 42219..42237  
 OTHER INFORMATION: 20-841-149.mis complement  
 NAME/KEY: primer\_bind  
 LOCATION: 45423..45441  
 OTHER INFORMATION: 20-842-115.mis  
 NAME/KEY: primer\_bind  
 LOCATION: 45443..45461  
 OTHER INFORMATION: 20-842-115.mis complement  
 NAME/KEY: primer\_bind  
 LOCATION: 77039..77057  
 OTHER INFORMATION: 20-853-415.mis  
 NAME/KEY: primer\_bind  
 LOCATION: 77059..77077  
 OTHER INFORMATION: 20-853-415.mis complement  
 NAME/KEY: misc\_binding  
 LOCATION: 1227..1251  
 OTHER INFORMATION: 20-828-311.probe  
 NAME/KEY: misc\_binding  
 LOCATION: 12315..12359  
 OTHER INFORMATION: 17-42-319.probe  
 NAME/KEY: misc\_binding  
 LOCATION: 15229..15253  
 OTHER INFORMATION: 17-41-250.probe  
 NAME/KEY: misc\_binding  
 LOCATION: 42206..42230  
 OTHER INFORMATION: 20-841-149.probe  
 NAME/KEY: misc\_binding  
 LOCATION: 45430..45454  
 OTHER INFORMATION: 20-842-115.probe  
 NAME/KEY: misc\_binding  
 LOCATION: 77046..77070  
 OTHER INFORMATION: 20-853-415.probe  
 US-09-750-580-1

Query Match 2.84; Score 169.8; DB 4; Length 81001;  
 Best Local Similarity 67.74; Pred. No. 1e-28;  
 Matches 270; Conservative 0; Mismatches 122; Indels 7; Gaps 2;  
 QY 3267 AATTTTAAAAAATAGATGATGCTGCTCATGCTGTATCCAAACATTGGG 3326

DB 38438 AATATTAATAAGAAATGCGCAGGTGCGGTGCTCATTCCTGTAATCTTAGCACTTTGGG 38379  
 QY 3327 AAGCGCGTCCGAGAGATGCTGAGTCACAGCTTTGAGACAGTACAGGCAACACAGC 3386  
 DB 38378 AGCTGAGGCGGGGTGATGCTTGAGCCAGAGATTATCATGACAGCTTAGGCAACGCT 38319  
 QY 3387 AAGACCCATATCTAAAAAACAACAAACAAATTTCTGGGTGTGTTGCTGAC 3446  
 DB 38318 GAAACCCGCTCTCT-----ACAAAAACAACAAATTTAGCTGGGCAATGCTGACGCTC 38265  
 QY 3447 CTGTATGCAAGCTACAGAGAGTGCAGCAAGCATCTTGAACCCAGAGAGCTTGA 3506  
 DB 38264 CTGTATCCAGCTATTGTGGAGGCTGAGGCAAGCTTCACTTGAACCTGGAGGACAA 38205  
 QY 3507 GCGTCAGTATCTGATGAAGCGGCTGCTACCT-CACTGCGGTGACAGTGCAGAACCT 3565  
 DB 38204 GCGTCAGTATGCAAGATTTGAGCCAGCTGCACTGCGGTGACAGTGCAGAACCT 38145  
 QY 3566 GCTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 3625  
 DB 38144 GCTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 38085  
 QY 3626 AAAACACACACATGAGATGTTGCAATTTGATTT 3664  
 DB 38084 ATACACCCACACATGAGCTTAATGACATTAATTAATTT 38046

RESULT 11  
 US-09-780-049-18/c  
 ; Sequence 18, Application US/09780049  
 ; Patent No. 6465250  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brett P. Monia  
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF PROTEIN PHOSPHATASE 2 CATALYTIC SUBUNIT  
 ; FILE REFERENCES: RTS-0134  
 ; CURRENT APPLICATION NUMBER: US/09/780,049  
 ; NUMBER OF SEQ ID NOS: 96  
 ; SEQ ID NO 18  
 ; LENGTH: 40000  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURES:  
 US-09-780-049-18

Query Match 2.84; Score 169.4; DB 4; Length 40000;  
 Best Local Similarity 71.84; Pred. No. 9.7e-29;  
 Matches 249; Conservative 0; Mismatches 96; Indels 2; Gaps 2;  
 QY 3257 TAAACCTTTAAATTTTAAAAAATAGATGATGCTGCTCATGCTGTATCCCA 3316  
 DB 12284 TAAATTTTAAAAATATATGTAAGGCTGGGTGAGTGGCTCAAGCTGTATCCCA 12225  
 QY 3317 ACACTTGGGAAGCCGGATGGGAGATAGCTTGAAGTCAAGCTTGAACAGCTGACG 3376  
 DB 12224 GACTTGGGAGGCGCGAGGAGGATGATGAGTCAAGGATCAAGGACAGCTGAG 12165  
 QY 3377 GCACACAGCAGACCCCATATCTMAAAAAACAAACAAATTAATCTGGGTATGG 3436  
 DB 12164 CCAAGTGTGTAACCCCATCTCTACCAAAA-AAAAAATACAAAAATTTCTGGGCAATGG 12106  
 QY 3437 TTGGCTCACTGATGATGACAGTACAGAGAGTGAAGGAGAGATCACTTGAAGCC 3496  
 DB 12105 TAGCCCTGCTGTGGTCCAGTATCTAGAGGCTGAGGAGAGAAATGCTTGAACCC 12046  
 QY 3497 AGAGAGTGAAGCTGACATGATCAATGAACGCTGCTACCT-CACTTGGGTGAAGAT 3555  
 DB 12045 AGAGAGTGAAGCTGACATGATCAATGAACGCTGCTACCT-CACTTGGGTGAAGAT 11986  
 QY 3556 GCAGAGAGCTGTCTCAAAATTAATTAATTAATTAATTAATTAATTTTAA 3602

11985 GCAAACTCCATCTCAAAAATAAAAATAAAAAGTAAAAA 11939

```

RESULT 12
US-09-780-172-18/c
; Sequence 18, Application US/09780172
; Patent No. 6607916
; GENERAL INFORMATION:
; APPLICANT: Robert McKay
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF CASEIN KINASE 2-ALPHA EXPRESSION
; FILE REFERENCE: PRS-0159
; CURRENT APPLICATION NUMBER: US/09/780,172
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 18
; LENGTH: 63000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-780-172-18

```

; LOCATION: (1) ... (111282)  
 ; OTHER INFORMATION: n = A, T, C or G  
 US-09-754-250-3





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 27, 2004, 07:46:34 ; Search time 2478 Seconds

(without alignments)  
12182.041 Million cell updates/sec

Title: US-10-030-294-2

Perfect score: 5960

Sequence: 1 cagcgttcaggagatgctg.....tttgcgtatattgcgac 5960

Scoring table: IDENTITY\_NUC

Searched: 3337386 seqs, 2532474682 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09D\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09E\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09F\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09G\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US09H\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US09I\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US09J\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US09K\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US09L\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US09M\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	196	3.3	304905	Sequence 1, Appl1
2	187.6	3.1	8133	Sequence 10, Appl1
3	186	3.1	18861	Sequence 513, App
4	185.6	3.1	611	Sequence 181221
5	185.6	3.1	611	Sequence 115440
6	185.4	3.1	3287	Sequence 115440
7	185.4	3.1	3287	Sequence 115440
8	182.6	3.1	47963	Sequence 808, App
9	182.4	3.1	736	Sequence 17813, A
10	182.4	3.1	736	Sequence 17813, A
11	182.2	3.1	36651	Sequence 3, Appl1
12	182.2	3.1	36651	Sequence 3, Appl1
13	181.8	3.1	50000	Sequence 8, Appl1
14	181.8	3.1	50000	Sequence 8, Appl1

15	181	3.0	58985	10	US-09-901-152-3	Sequence 3, Appl1
16	181	3.0	143601	10	US-09-855-824-3	Sequence 3, Appl1
17	180.8	3.0	2214	13	US-10-027-632-110324	Sequence 110324
18	180.8	3.0	2214	16	US-10-027-632-110324	Sequence 110324
19	180.8	3.0	198522	13	US-10-087-192-244	Sequence 244, App
20	180	3.0	94720	12	US-10-052-482-160	Sequence 160, App
21	179.8	3.0	108316	15	US-10-292-798-1789	Sequence 1789, App
22	179.8	3.0	108316	16	US-10-017-161-2143	Sequence 2143, App
23	179.6	3.0	51657	16	US-10-057-475B-10475	Sequence 10475, A
24	179.6	3.0	51657	16	US-10-154-844B-10475	Sequence 10475, A
25	179.6	3.0	65608	9	US-09-962-436-292	Sequence 292, App
26	179.6	3.0	65608	9	US-09-962-833-119	Sequence 119, App
27	179.6	3.0	65608	9	US-09-954-531-180	Sequence 180, App
28	179.2	3.0	812	13	US-10-027-632-170028	Sequence 170028
29	179.2	3.0	812	16	US-10-027-632-170028	Sequence 170028
30	179.2	3.0	94330	13	US-10-087-192-1222	Sequence 1222, App
31	178.6	3.0	186957	15	US-10-185-770-3	Sequence 3, Appl1
32	178.6	3.0	201986	13	US-10-087-192-874	Sequence 874, App
33	178.4	3.0	567	13	US-10-027-632-136457	Sequence 136457
34	178.4	3.0	567	16	US-10-027-632-136457	Sequence 136457
35	178.2	3.0	906	15	US-10-198-846-7052	Sequence 7052, App
36	178.2	3.0	25423	9	US-09-764-855-193	Sequence 193, App
37	178.2	3.0	25423	15	US-10-072-349-193	Sequence 193, App
38	178.2	3.0	25424	9	US-09-764-855-194	Sequence 194, App
39	178.2	3.0	25424	15	US-10-072-349-194	Sequence 194, App
40	178	3.0	48001	17	US-10-304-107-4	Sequence 4, Appl1
41	177.8	3.0	56826	17	US-10-322-281-664	Sequence 664, App
42	177.6	3.0	17588	15	US-10-017-161-1603	Sequence 1603, App
43	177.6	3.0	96256	17	US-10-292-798-1277	Sequence 1277, App
44	177.6	3.0	96256	17	US-10-775-169-352	Sequence 352, App
45	177.2	3.0	9566	9	US-09-764-864-1768	Sequence 1768, App

## ALIGNMENTS

RESULT 1

US-10-271-416-1/c

Sequence 1, Application US/10271416

Publication No. US20040043021A1

GENERAL INFORMATION:

APPLICANT: Keith, Tim

APPLICANT: Little, Randall D.

APPLICANT: Van Berdeewegh, Paul

APPLICANT: Dupuis, Joseph

APPLICANT: Del Maestro, Richard G.

TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES

FILE REFERENCE: 2976-4045

CURRENT APPLICATION NUMBER: US/10/271,416

CURRENT FILING DATE: 2002-10-11

PRIOR APPLICATION NUMBER: 60/328,424

PRIOR FILING DATE: 2001-10-11

NUMBER OF SEQ ID NOS: 9

SOFTWARE: FASTSEQ For Windows Version 4.0

SEQ ID NO 1

LENGTH: 304905

TYPE: DNA

ORGANISM: Homo sapien

Query Match 3.3%; Score 196, DB 13; Length 304905;

Best Local Similarity 73.3%; Pred. No. 4.2e-30;

Matches 264; Conservative 0; Mismatches 95; Indels 1; Gaps 1;

DB 156005

3279

3339

155945

GAGATGTCCTTAACCCAGGAGTTTGAACCAAGCTGGGCAACATGATGTAACCCATGT

155886

Query Match	Similarity	Score	187.6	DB 13	Length 8133
Best Local Similarity		74.6%	Pred. No. 3,4e-29		
Matches	249	Conservative	0	Mismatches	84
				Indels	1
				Gaps	1
Qy	3264	TTTAAATTTTAAAAAATAGATGAGTGGTGGCTGCTGCTGTAATCCCAACGTT	3323		
Db	1057	TTTAAAGAGAAAACAGTACGCTGGGACGCTGGCTCAAGCTGTGATGACAGCTTT	1166		
Qy	3324	GGGAAGCGGTGGGAGATACCTGATGCTCAAGCTTTGAGAACGACTGAGGCAAC	3385		
Db	1117	GAGAGCGAAGCGGAGGATCATTTGAGCGAGATTGAGACAGACCTGGGCAAGAT	1176		
Qy	3384	AGCAAGACCCCATATCTTAAAAAACAACAAAACAATTACCTGGTATGTTGTGCT	3443		
Db	1177	AGTGAAGCCCTGCTCTTCAAAAAAATTAAAAATTAGCCGAGGTGTGTGTGCA	1236		
Qy	3444	CACCTGATGCTCAAGTACACAGAAAGCTGAGCGAGAGATCACTTGAGCCACAGGT	3503		
Db	1237	CACCTGATGCTCAAGTACCTGGGAGGCTGAGGCAAGAGATCACTTGAGCTGGAAGTT	1296		
Qy	3504	TGAGGCTGCAAGTATCATGAACGGCTGCTCAACT-CAGTGTGGGTGAACATGACGA	3562		
Db	1297	GGAAGCTGCAAGTATGATTTGACCACTTGGCTTCAAGCTGGGCAACGAGTGAAC	1356		
Qy	3563	GCTGCTCAAAAAATTAATTAATTAATTAATTAATTAAC	3596		
Db	1357	CCTGCTCAAAATTTAAAAAATTAATTAATTAATTAAC	1390		

RESULT 3  
 US-09-984-429-513  
 Sequence 513 Application US/09984429  
 Publication No. US20040010132A1  
 GENERAL INFORMATION:  
 APPLICANT: Rohen et al.  
 TITLE OF INVENTION: 53 Human Secreted Proteins

Query Match	3.1%	Score 186;	DB 11;	Length 18861;
Best Local Similarity	76.1%	Pred. No. 1.2e-28;		
Matches 242;	Conservative	0;	Mismatches 75;	Indels 1;
				Gaps 1;
Qy	3265	TTAAATTTTAAAAAATAAGATGAGTGTGTCTCAGACCCCTGATATCCCAACTTGG	3324	
Db	5557	TAAAGTGAATAAATAATGACTAGGTGCACAGCGCTCAATGTCTATATCCACAGCCTTTG	5616	
Qy	3325	GGAGCCCGGCTCGGAGGAGTACGTTGAGTCCAGCAGTTTGAGACACGATCAGGCAACCA	3384	
Db	5617	GGAGCCCGGAGGAGGAGATCACTTGAGCTCAGAGAGTTTGAGACACGCTGAGCAACATA	5676	
Qy	3385	GCAAGACCCCATATGTCTAAATAAACAACAACAACAATAATACCTGGTATGAGTTGGTCT	3444	
Db	5677	GCAAGCTTCACTCTACTTAAATAAAGAAAGTATTAATAATTTAGCAGAGCCTGGTGGTGCAC	5736	
Qy	3445	ACCTGTATGCCAGCTACACAGAGGCTGAGGCAGAAAGATCACTTGAGCCCAAGAGGTT	3504	
Db	5737	ACCTGTATGCCAGCTACACAAAGGCTGAGGCAGAGGATCACTTGAGTCTGGAGAGGTT	5796	
Qy	3505	GAGGCTGCAGTATCATCATGAACGCCCTGCTCACTCACTGCTGGGTGAACGTGCAAGAG	3563	
Db	5797	GAGGCTGCAGTATCATCATGATCCGACACACGACTCACTCACTGCTGGGTGAACGTACGCT	5856	
Qy	3564	CTGCTCAAAAAATAATAA 3581		
Db	5857	CTGCTCCAAAAGACAA 5874		

RESULT 4  
 US-10-027-632-181221/c  
 ; Sequence 181221, Application US/10027632  
 ; Publication No. US20020198371A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, David G.  
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
 ; FILE REFERENCE: 108827.129  
 ; CURRENT APPLICATION NUMBER: US/10/027.632  
 ; PRIOR FILING DATE: 2002-04-10  
 ; PRIOR APPLICATION NUMBER: US 60/218,006  
 ; PRIOR FILING DATE: 2000-07-12  
 ; PRIOR APPLICATION NUMBER: US 60/198,676  
 ; PRIOR FILING DATE: 2000-04-20

```

; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 181221
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-181221

```

```

Query Match      3.1%; Score 185.6; DB 13; Length 611;
Best Local Similarity 71.6%; Pred. No. 2.1e-29;
Matches 272; Conservative 0; Mismatches 104; Indels 4; Gaps 2;

```

```

QY 3251 AGCTATTAACCTTTTAAATTTTAAATAATGATGAGTGTGTGCTCATGCTGTA 3310
DB 410 AGGTTTCACCTGTTTAAAAAAGAAACAGGGGCTGACATGCTGCTCATGCTGTA 351
QY 3311 ATCCCAACCTTTGGGAGCCGGGTGGGAGATAGCTTGAATCCAGCATTTGACCA 3370
DB 350 ATCCAGACCTTTGGGAGCCAGGAGCATGATGCTTGAATCCAGCATTTGACCA 291
QY 3371 GTGAGGCAACACAGACAGACCCCATCT---AAAAAACAACAACAATAATAC 3427
DB 290 GCTGAAACAAATAGGAGACCTGCTCTACAAAATAAAAAAAAAAAAAAAAAATG 231
QY 3428 TGGGTATGTTGTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3487
DB 230 TGGGGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 171
QY 3488 CTGAGCCCAAGAGGTTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3546
DB 170 CTGAGCCCTGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 111
QY 3547 GGTACAGTGCAGAGAGCTGTCTCAAAATAATTAATAATAATACTTTAAAAA 3606
DB 110 GGTACAGAGACAGATCTGTCTCAAAATAATTAATAATAATAATAATAATA 51
QY 3607 CAAAATTAATTAATTTTA 3626
DB 50 TAGATATGTTCCCTTTCTA 31

```

```

RESULT 5
US-10-027-632-181221/c
; Sequence 181221, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358

```

```

; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 181221
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-181221

```

```

Query Match      3.1%; Score 185.6; DB 16; Length 611;
Best Local Similarity 71.6%; Pred. No. 2.1e-29;
Matches 272; Conservative 0; Mismatches 104; Indels 4; Gaps 2;

```

```

QY 3251 AGCTATTAACCTTTTAAATTTTAAATAATGATGAGTGTGTGCTCATGCTGTA 3310
DB 410 AGGTTTCACCTGTTTAAAAAAGAAACAGGGGCTGACATGCTGCTCATGCTGTA 351
QY 3311 ATCCCAACCTTTGGGAGCCGGGTGGGAGATAGCTTGAATCCAGCATTTGACCA 3370
DB 350 ATCCAGACCTTTGGGAGCCAGGAGCATGATGCTTGAATCCAGCATTTGACCA 291
QY 3371 GTGAGGCAACACAGACAGACCCCATCT---AAAAAACAACAACAATAATAC 3427
DB 290 GCTGAAACAAATAGGAGACCTGCTCTACAAAATAAAAAAAAAAAAAAAAAATG 231
QY 3428 TGGGTATGTTGTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3487
DB 230 TGGGGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 171
QY 3488 CTGAGCCCAAGAGGTTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3546
DB 170 CTGAGCCCTGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 111
QY 3547 GGTACAGTGCAGAGAGCTGTCTCAAAATAATTAATAATAATACTTTAAAAA 3606
DB 110 GGTACAGAGACAGATCTGTCTCAAAATAATTAATAATAATAATAATAATA 51
QY 3607 CAAAATTAATTAATTTTA 3626
DB 50 TAGATATGTTCCCTTTCTA 31

```

```

RESULT 6
US-10-027-632-115440/c
; Sequence 115440, Application US/10027632
; Publication No. US2002019837A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115440
; LENGTH: 3287

```











GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: September 27, 2004, 07:46:33 ; Search time 2018 seconds  
(without alignments)

12423.582 Million cell updates/sec

Title: US-10-030-294-2

Perfect score: 5960  
Sequence: 1 cagctgttcggagatctcg.....tttcgttcgatttcgac 5960

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 212409941 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

N\_Geneseq\_29Jan04:\*  
1: geneseq1980s:\*  
2: geneseq1990s:\*  
3: geneseq2000s:\*  
4: geneseq2001as:\*  
5: geneseq2001bs:\*  
6: geneseq2002s:\*  
7: geneseq2003as:\*  
8: geneseq2003bs:\*  
9: geneseq2003cs:\*  
10: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5960	100.0	5960	5	AAF25905 Human tum
2	5681.6	95.3	13940	5	AAF25915 Human tum
3	5676	95.2	5676	5	AAF25904 Human tum
4	5676	95.2	5676	5	AAF25907 Human tum
5	5676	95.2	5676	5	AAF25906 Human tum
6	193.4	3.2	118384	7	ABX56555 Human aut
7	189.4	3.2	118384	7	ABX56555 Human aut
8	187.6	3.1	8133	6	ABN99663 Human clu
9	185.4	3.1	47939	7	AAV252898 Human lwe
10	183	3.1	335199	9	ADC24703 Human wll
11	182.2	3.1	36651	9	ADC24703 Human wll
12	181.8	3.1	49939	8	ADH16928 Human dyc
13	181	3.0	58985	8	ABZ59738 Human sec
14	181	3.0	143601	7	ABZ22654 Human epi
15	180	3.0	26329	4	AAH420654 Human STR
16	180	3.0	94720	9	ADH272982 Human STR
17	180	3.0	94720	9	ADH272982 Human STR
18	179.8	3.0	108316	9	ADH27336 Human GPC
19	179.6	3.0	65608	6	AB162910 Breast ca
20	179.6	3.0	65608	6	AB162910 Breast ca
21	179.6	3.0	65608	6	AB162910 Breast ca
22	178.5	3.0	186957	7	AB167668 Geophagnu
23	178.2	3.0	630	4	AA15941 Human bre

24	178.2	3.0	7726	5	ABA21079 Human ner
25	178.2	3.0	25423	4	AAK90279 Human dig
26	178.2	3.0	25423	4	AAI57656 Human col
27	178.2	3.0	25423	6	ABG99833 Genomic D
28	178.2	3.0	25423	6	ADH92986 Human col
29	178.2	3.0	25424	4	AAK90280 Human col
30	178.2	3.0	25424	4	AAI57657 Human col
31	178.2	3.0	25424	6	ABG99834 Genomic D
32	178.2	3.0	25424	6	ADH92987 Human col
33	177.8	3.0	16310	3	AAI34964 Human low
34	177.8	3.0	16310	3	AAI34964 Human low
35	177.8	3.0	16310	3	AAI34965 Human low
36	177.8	3.0	17634	3	AAI34965 Human low
37	177.8	3.0	17634	3	AAI34965 Human low
38	177.8	3.0	17634	7	ABZ21087 Human nuc
39	177.6	3.0	17588	9	ADH86824 Human GPC
40	177.2	3.0	9566	4	AAI37121 Human mus
41	177.2	3.0	9566	4	AAI37121 Human mus
42	177.2	3.0	9566	5	ABX74143 Human ner
43	177.2	3.0	9566	7	ABX74143 Human nov
44	177.2	3.0	9566	7	ABX60109 CDNA enco
45	177.2	3.0	183610	7	ACF62736 Cancer ba

## ALIGNMENTS

RESULT 1  
AAF25905  
ID AAF25905 standard; DNA: 5960 BP.  
AC AAF25905;  
DT 19-APR-2001 (first entry)  
DE Human tumor suppressor gene p51 promoter associated DNA SEQ ID 2.  
KW Tumor suppressor; p51, cell death, cell proliferation; cancer;  
KW cytoskeletal; gene therapy; screening; ds.  
XX Homo sapiens.  
OS  
XX  
XX  
XX WO200100818-A1.  
XX  
XX 04-JAN-2001.  
XX  
XX 28-JUN-2000; 2000WO-JP004261.  
XX  
XX 29-JUN-1999; 98JP-00183195.  
XX  
XX (NIPK.) NIPPON KAYAKU KK.  
XX (SAKAI.) SAKAI T.  
XX Sakai T, Kagaya S, Sato T, Sukenaga Y, Fujii H;  
XX WPI, 2001-112452/12.  
XX  
XX New DNA sequences encoding the human p51 promoter domain for screening  
XX for compounds that modify p51 promoter activity and for diagnosing and  
XX treating cancer.  
XX  
XX Claim 1(4); Page 35-36; 60pp; Japanese.  
XX  
XX This invention describes novel DNA sequences (1) encoding the human p51  
XX promoter domain (which may include the 5'-untranslated sequence) or  
XX derived from it by addition, deletion and/or substitution of one or more  
XX bases. The invention also describes (1) expression plasmids including (1)  
XX (2) host cells transformed by (1); (3) DNA probes binding to all or  
XX part of (1); (4) cloning (1) using (3); (5) DNA (and RNA corresponding to  
XX (1)) inhibiting the expression of p51, which includes the antisense  
XX sequence of (1); (6) screening compounds for their ability to modify p51  
XX promoter activity, by observing their effect on cells transformed by (1);  
XX (7) compounds identified by (6); and (8) drug compositions containing

CC (7). The products of the invention have cytostatic activity and can be  
 CC used for gene therapy. (1) is used to screen for compounds that modify  
 CC p51 promoter activity. (1) and RNA corresponding to it are used to  
 CC inhibit the expression of p51. (1) and the identified compounds are used  
 CC for the diagnosis and treatment of cancer

XX Sequence 5960 BP, 1858 A, 1103 C, 1175 G, 1824 T, 0 U, 0 Other,

Query Match 100.0%; Score 5960; DB 5; Length 5960;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;

Matches 5960; Conservative 0; Indels 0; Gaps 0;

QY 1 CAGCTGTTCAAGGATGCTGGAAGAAAGCCCAACCTGCTTGAACATCGGCT 60  
 DB 1 CAGCTGTTCAAGGATGCTGGAAGAAAGCCCAACCTGCTTGAACATCGGCT 60  
 QY 61 GACTTGGAGGATGATGAGGTTGTCTGTTAAAGAACTGCAACTCTTCTGCCCAAT 120  
 DB 61 GACTTGGAGGATGATGAGGTTGTCTGTTAAAGAACTGCAACTCTTCTGCCCAAT 120  
 QY 121 TGGCTCTGTTCCCTTGAAGCCCTCTTCTTGGGACCTCCCTTAAGGACATCTTCTG 180  
 DB 121 TGGCTCTGTTCCCTTGAAGCCCTCTTCTTGGGACCTCCCTTAAGGACATCTTCTG 180  
 QY 181 ACATTAACCTAATTAATGTTTATTTGATGAATTTCACTGAGAGAGATGAG 240  
 DB 181 ACATTAACCTAATTAATGTTTATTTGATGAATTTCACTGAGAGAGATGAG 240  
 QY 241 GTCATTAAGAAAGCAATGAGGCTAAGGTTGCAATGCACTGCTTTTCAATTA 300  
 DB 241 GTCATTAAGAAAGCAATGAGGCTAAGGTTGCAATGCACTGCTTTTCAATTA 300  
 QY 301 AGCTATTCGAATACATTAAGGTTTACTTAAGTTCAGGCCCAAGTTTCTCAAT 360  
 DB 301 AGCTATTCGAATACATTAAGGTTTACTTAAGTTCAGGCCCAAGTTTCTCAAT 360  
 QY 361 GTCAGACTGAGCAATTAAGGTTGCAAGTTGGAAGATTAAGAGATCTCTCA 420  
 DB 361 GTCAGACTGAGCAATTAAGGTTGCAAGTTGGAAGATTAAGAGATCTCTCA 420  
 QY 421 CAGAACTTGGCTTCACTTCTAATAAATGAGATGTTGTTGAGTCAAGCTATGA 480  
 DB 421 CAGAACTTGGCTTCACTTCTAATAAATGAGATGTTGTTGAGTCAAGCTATGA 480  
 QY 481 CAGAACTTGGCTTCACTTCTAATAAATGAGATGTTGTTGAGTCAAGCTATGA 480  
 DB 481 CAGAACTTGGCTTCACTTCTAATAAATGAGATGTTGTTGAGTCAAGCTATGA 480  
 QY 541 TGGTGAATTTACTGCACTCTCTCTCTGAAAGTCTTCTCTCTCTCTCAATA 600  
 DB 541 TGGTGAATTTACTGCACTCTCTCTCTGAAAGTCTTCTCTCTCTCTCAATA 600  
 QY 601 AATGCAAGATGACCTGGAAGGCTAGACCTGAGGTTCAATACCTGACACAAAGAT 660  
 DB 601 AATGCAAGATGACCTGGAAGGCTAGACCTGAGGTTCAATACCTGACACAAAGAT 660  
 QY 661 AATGCAAGATGACCTGGAAGGCTAGACCTGAGGTTCAATACCTGACACAAAGAT 660  
 DB 661 AATGCAAGATGACCTGGAAGGCTAGACCTGAGGTTCAATACCTGACACAAAGAT 660  
 QY 721 TCACTTCTCTGATCTGATAGTCAAGCTGCAAGCTCTAAGGAATGCAATGAT 720  
 DB 721 TCACTTCTCTGATCTGATAGTCAAGCTGCAAGCTCTAAGGAATGCAATGAT 720  
 QY 781 TCCCTTGGTTAAGGAACTTTTCTTCTGAAAGCTCTGCTCAATGATGAGG 840  
 DB 781 TCCCTTGGTTAAGGAACTTTTCTTCTGAAAGCTCTGCTCAATGATGAGG 840  
 QY 841 CCACTGATCCAGATGATGAGGCTCTCTGAGCAACATGCACTTTTCTTGGAAAT 900  
 DB 841 CCACTGATCCAGATGATGAGGCTCTCTGAGCAACATGCACTTTTCTTGGAAAT 900  
 QY 901 TGAATCTTAAAGCTGAATAGCTGAAGTTCAAAAAGCTGTTGATCTGACTAGCCTAC 960

DB 901 TGAATCTTAAAGCTGAATAGCTGAAGTTCAAAAAGCTGTTGATCTGACTAGCCTAC 960  
 QY 961 AGTGGCTTTCGAAAGTACCTGATTCATATTTCTTAAGTCCCTGAATTAATTAAT 1020  
 DB 961 AGTGGCTTTCGAAAGTACCTGATTCATATTTCTTAAGTCCCTGAATTAATTAAT 1020  
 QY 1021 CCTGGTACAGCCCTTCTGAGATGATGATGTTTCTTCCAACTGCTCTATAGCTGT 1080  
 DB 1021 CCTGGTACAGCCCTTCTGAGATGATGATGTTTCTTCCAACTGCTCTATAGCTGT 1080  
 QY 1081 GAATTTTCAATTTCTTTTCAATCAATTTTCAATGTTTGTGTTGTTGTTT 1140  
 DB 1081 GAATTTTCAATTTCTTTTCAATCAATTTTCAATGTTTGTGTTGTTGTTT 1140  
 QY 1141 GCTTTAGTACGAGATCAATGTTCTGTTGTTTAAACCAAGGAATCTGATTAACA 1200  
 DB 1141 GCTTTAGTACGAGATCAATGTTCTGTTGTTTAAACCAAGGAATCTGATTAACA 1200  
 QY 1201 TCTTCCCTTTAAATTAATGATCTAAGGCTCAAGGATGAGTACTGCTGAGG 1260  
 DB 1201 TCTTCCCTTTAAATTAATGATCTAAGGCTCAAGGATGAGTACTGCTGAGG 1260  
 QY 1261 TCTGAGATTAAGTATGATCAAGCTCTGATTAACCCAGGTTAGCCACTGCTTAC 1320  
 DB 1261 TCTGAGATTAAGTATGATCAAGCTCTGATTAACCCAGGTTAGCCACTGCTTAC 1320  
 QY 1321 AACATTTGCTCTCTCTGATGATTAAGGATCTGGAAGAAAGATTAATTTG 1380  
 DB 1321 AACATTTGCTCTCTCTGATGATTAAGGATCTGGAAGAAAGATTAATTTG 1380  
 QY 1381 CAAAGCTTCAAGGATCAAGCAAGCTTGGGATTTGATGATTAATCTTAA 1440  
 DB 1381 CAAAGCTTCAAGGATCAAGCAAGCTTGGGATTTGATGATTAATCTTAA 1440  
 QY 1441 AGTTTGTGATCTTCTGCAATTAATGTTGATTAAGCAAGATTTCTTAATAGCT 1500  
 DB 1441 AGTTTGTGATCTTCTGCAATTAATGTTGATTAAGCAAGATTTCTTAATAGCT 1500  
 QY 1501 AATCAACCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCA 1560  
 DB 1501 AATCAACCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCA 1560  
 QY 1561 AGTTGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCA 1620  
 DB 1561 AGTTGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCA 1620  
 QY 1621 AAGAGAGATAGTACCAATGCTGCAATTTATGATATCTAGATTTATGTTTCA 1680  
 DB 1621 AAGAGAGATAGTACCAATGCTGCAATTTATGATATCTAGATTTATGTTTCA 1680  
 QY 1681 CATAGTACCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCA 1740  
 DB 1681 CATAGTACCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCA 1740  
 QY 1741 AATAAATTTAGCTCTTCAAGTACCTGATCAAGTAACTAGGCTACACATGAC 1800  
 DB 1741 AATAAATTTAGCTCTTCAAGTACCTGATCAAGTAACTAGGCTACACATGAC 1800  
 QY 1801 GATGCAATTAAGGATTTCTTAAACAGAAAGTCTATTGGAAGAAATGTTCTAG 1860  
 DB 1801 GATGCAATTAAGGATTTCTTAAACAGAAAGTCTATTGGAAGAAATGTTCTAG 1860  
 QY 1861 AAAATATACATTAATCTAATTAAGGATCAAGGATCAAGGATCAAGGATCA 1920  
 DB 1861 AAAATATACATTAATCTAATTAAGGATCAAGGATCAAGGATCAAGGATCA 1920  
 QY 1921 TGAAGGAGATGATGATCAAGGATGATGATGATGATGATGATGATGATGAT 1980  
 DB 1921 TGAAGGAGATGATGATCAAGGATGATGATGATGATGATGATGATGATGAT 1980  
 QY 1981 ATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040



Oy	4201	CACACATTTAGCCGCTCTGGGCTAGTAAAGAAAGAAATGTCAGATTTAAATAGT	4260
Dp	4201	CACGATGTTAGCGGCTCTGGGCTAGTAAAGAAAGAAATGTCAGATTTAAATAGT	4260
Oy	4261	TCTCCCTTCCATCTCTGGCTGAAGCACAATTAATTTTANGAACACATTTGAGT	4320
Dp	4261	TCTCCCTTCCATCTCTGGCTGAAGCACAATTAATTTTANGAACACATTTGAGT	4320
Oy	4321	TGAGATTACTTACAGGGAATGTCCAAATTTCTCGAAGAGGGCTTTAGATGTCTCAAC	4380
Dp	4321	TGAGATTACTTACAGGGAATGTCCAAATTTCTCGAAGAGGGCTTTAGATGTCTCAAC	4380
Oy	4381	TTTGACATCTACAGATGACCTATTATACAGAGTGTCTCGTACAGAGGGGTGAAGGGA	4440
Dp	4381	TTTGACATCTACAGATGACCTATTATACAGAGTGTCTCGTACAGAGGGGTGAAGGGA	4440
Oy	4441	AAGTGTAACTCACATGTCTTAGTACCGTTAGATACACAGAGTGGTTTTTTTTTCCCTG	4500
Dp	4441	AAGTGTAACTCACATGTCTTAGTACCGTTAGATACACAGAGTGGTTTTTTTTTCCCTG	4500
Oy	4501	TTGAGATCTATCCCTAACGAGCTTCTGATCATATTTCAATTCACAAATCCACAA	4560
Dp	4501	TTGAGATCTATCCCTAACGAGCTTCTGATCATATTTCAATTCACAAATCCACAA	4560
Oy	4561	ACCAGAGAAAGTTTACAGGCCATCTCAGAAAGAAATTAATTTGTCGTACACTT	4620
Dp	4561	ACCAGAGAAAGTTTACAGGCCATCTCAGAAAGAAATTAATTTGTCGTACACTT	4620
Oy	4621	TCTGATATTTACACTGATTTGGGAATATATGAACAATTTATGCTTCTTCCAGATAG	4680
Dp	4621	TCTGATATTTACACTGATTTGGGAATATATGAACAATTTATGCTTCTTCCAGATAG	4680
Oy	4681	GTCAGATCAAGCAAAACCAAAACACGAAAACTGTAAACATATTAAGAAATTAAGTGAG	4740
Dp	4681	GTCAGATCAAGCAAAACCAAAACACGAAAACTGTAAACATATTAAGAAATTAAGTGAG	4740
Oy	4741	CCGATGAGAGATTTAAAACTAGAAATTTTTTTTAAACAGGCAATTTGAAATAATT	4800
Dp	4741	CCGATGAGAGATTTAAAACTAGAAATTTTTTTTAAACAGGCAATTTGAAATAATT	4800
Oy	4801	GTGCACTTCAGAAATTTCTACATATATATATTTTCCAAATTTAATCTTTAGAGAAA	4860
Dp	4801	GTGCACTTCAGAAATTTCTACATATATATATTTTCCAAATTTAATCTTTAGAGAAA	4860
Oy	4861	TTACTATATTAATATGTAAAGTACATGTCAGTGTGAGTGAATTTAACTCAATA	4920
Dp	4861	TTACTATATTAATATGTAAAGTACATGTCAGTGTGAGTGAATTTAACTCAATA	4920
Oy	4921	AGCTAATTTCTTTAATTCGGGTCCAGCAAAAGCTCTAAGCGGATGTGAAGGCATATCT	4980
Dp	4921	AGCTAATTTCTTTAATTCGGGTCCAGCAAAAGCTCTAAGCGGATGTGAAGGCATATCT	4980
Oy	4981	CTTTCTCTTAGCTAGAGAAAGTGAAGTTCTAAGTTAAATATATCAAGCAATTTCCCT	5040
Dp	4981	CTTTCTCTTAGCTAGAGAAAGTGAAGTTCTAAGTTAAATATATCAAGCAATTTCCCT	5040
Oy	5041	GTCCTTGTCTTAGATATGTGACCAACACAGCGGTGTGGCTGAAGGGGAACTGAAGGG	5100
Dp	5041	GTCCTTGTCTTAGATATGTGACCAACACAGCGGTGTGGCTGAAGGGGAACTGAAGGG	5100
Oy	5101	CGGGAAGGAGGAAATAGATGAAAAACAAACCAAACTTCCCTAAGCACTCT	5160
Dp	5101	CGGGAAGGAGGAAATAGATGAAAAACAAACCAAACTTCCCTAAGCACTCT	5160
Oy	5161	ACAAAACATTTTAGCCCCAGAAATAGTCAAGAAATCTCAAAATCAAAACAGATTCAGA	5220
Dp	5161	ACAAAACATTTTAGCCCCAGAAATAGTCAAGAAATCTCAAAATCAAAACAGATTCAGA	5220
Oy	5221	TACAAAGAAAGTTATATGATAGCTGAGAGAGGGGTGACACTCATGACCTCAATTCAGTTACA	5280
Dp	5221	TACAAAGAAAGTTATATGATAGCTGAGAGAGGGGTGACACTCATGACCTCAATTCAGTTACA	5280
Oy	5281	AAAGTCCAGGCTCTGAATTTAACTCTGATGCCATTCATGTCAGCATTCATACACACA	5340

Db		581	AAGTCACGCTCGTGAATTTAACTGTATGGCATTATGCCAGATCTCAATCCAGCA	5340
Oy		5341	GAAATAGAAGTTCAAGATGCCTCCAGCTCCAAATTGGCACACAAAGTGCTACTA	5400
Db		5341	GAAATCGAAGTTCAAGATGCTCCAGCTCCAAATTGGCACAAAGTGCTACTA	5400
Oy		5401	TACGTCAGACCTCTGAAAGCCGTGAGAGAGGAGAAACAACAGTGGAGATGCCCA	5460
Db		5401	TACGTCAGACCTCTGAAAGCCGTGAGAGAGGAGAAACAACAGTGGAGATGCCCA	5460
Oy		5461	GCTGTGTAAGAACGAGTGTATTATGAGTTTTAGTCATATGATCATCTCATTTGCTAAA	5520
Db		5461	GCTGTGTAAGAACGAGTGTATTATGAGTTTTAGTCATATGATCATCTCATTTGCTAAA	5520
Oy		5521	TCAAAGAAAGCTCCGCTCTTTGGAAATATGTATGAAAGAGAGAGTGCCTAACTCTA	5580
Db		5521	TCAAAGAAAGCTCCGCTCTTTGGAAATATGTATGAAAGAGAGAGTGCCTAACTCTA	5580
Oy		5581	TGTCGTATAGCATTTGACCCTATTGCTTTTAGCTCCGCTCTTATCTATATATAC	5640
Db		5581	TGTCGTATAGCATTTGACCCTATTGCTTTTAGCTCCGCTCTTATCTATATATAC	5640
Oy		5641	AGGTATTTGTATAATTTATATATATATTTGTTCTCGTTGTTGATA-TCAAAGCAGTTGAA	5700
Db		5641	AGGTATTTGTATAATTTATATATATATTTGTTCTCGTTGTTGATA-TCAAAGCAGTTGAA	5700
Oy		5701	GCAAAATGATTTTGGAACTTCACGCGTGGCCACCTACAGTACTGCCCCGACCTTACAT	5760
Db		5701	GCAAAATGATTTTGGAACTTCACGCGTGGCCACCTACAGTACTGCCCCGACCTTACAT	5760
Oy		5761	CCAGCGGTAGTTTGAATATATATATATATCTCTGCAAACTTAATGAAAGGCCGTGT	5820
Db		5761	CCAGCGGTAGTTTGAATATATATATATATCTCTGCAAACTTAATGAAAGGCCGTGT	5820
Oy		5821	ATTATGAATATGTCTACGCTGTGTACAAAGAACATTTCTCTTGTATTAGTCACACAGTGA	5880
Db		5821	ATTATGAATATGTCTACGCTGTGTACAAAGAACATTTCTCTTGTATTAGTCACACAGTGA	5880
Oy		5881	TATTAATTTTGGACCTTCTGTGTGACCTTAAGTGCTGTGTGACATATTTTCTGAAATGCTT	5940
Db		5881	TATTAATTTTGGACCTTCTGTGTGACCTTAAGTGCTGTGTGACATATTTTCTGAAATGCTT	5940
Oy		5941	TTTTGGTTGATATTTGGATC 5960	
Db		5941	TTTTGGTTGATATTTGGATC 5960	
<b>RESULT 2</b>				
AAF25915				
ID	AAF25915	strand:	DNA: 13940 BP.	
NC	AAF25915;			
XX	19-APR-2001 (first entry)			
DE	Human tumor suppressor gene p51 promoter DNA fragment SEQ ID 12.			
XX	Tumor suppressor; p51; cell death; cell proliferation; cancer;			
KM	cytostatic; gene therapy; screening; db.			
XX	Homo sapiens.			
OS	Homo sapiens.			
XX	MOZ00100818-A1.			
PV	04-JAN-2001.			
PD	28-JUN-2000; 2000MO-JP004261.			
XX	29-JUN-1999; 99JP-00183195.			
PR	(NIPK.) NIPON KAYAKU KK.			
XX	(SAKA.) SAKAI T.			
PA				

XX Sekai T, Kagaya S, Sato T, Sukemaga Y, Fujii H;  
 XX WPI; 2001-112452/12.

XX New DNA sequences encoding the human p51 promoter domain for screening  
 PT for compounds that modify p51 promoter activity and for diagnosing and  
 PT treating cancer.

XX Example 1; Page 49-57; 60pp; Japanese.

XX This invention describes novel DNA sequences (I) encoding the human p51  
 CC promoter domain (which may include the 5'-untranslated sequence) or  
 CC derived from it by addition, deletion and/or substitution of one or more  
 CC bases. The invention also describes (1) expression plasmids including (1)  
 CC ; (2) host cells transformed by (1); (3) DNA probes binding to a11 or  
 CC part of (1); (4) cloning (1) using (3); (5) DNA (and RNA corresponding to  
 CC 1c) inhibiting the expression of p51, which includes the antisense  
 CC sequence of (1); (6) screening compounds for their ability to modify p51  
 CC promoter activity, by observing their effect on cells transformed by (1);  
 CC (7) compounds identified by (6); and (8) drug compositions containing  
 CC (7). The products of the invention have cytotoxic activity and can be  
 CC used for gene therapy. (I) is used to screen for compounds that modify  
 CC p51 promoter activity. (I) and RNA corresponding to it are used to  
 CC inhibit the expression of p51. (1) and the identified compounds are used  
 CC for the diagnosis and treatment of cancer

XX Sequence 13940 BP; 3949 A; 2972 C; 3085 G; 3934 T; 0 U; 0 Other;

XX Query Match 95.3%; Score 5681.6; DB 5; Length 13940;

XX Best Local Similarity 99.3%; Pred. No. 0;

Matches 5705; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 2 AGCTGTTGAGGATGCTGCAAAAGAACCCCACTTCTTGTGCACTGGGTGTG 61  
 DB ATCTGTTGAGGATGCTGCAAAAGAACCCCACTTCTTGTGCACTGGGTGTG 110  
 QY 62 ACTTGGAGGGATGAGTGTGTTGTTAAAGAACTGCCAATCTCTTCTGCCAAT 121  
 DB ACTTGGAGGGATGAGTGTGTTGTTAAAGAACTGCCAATCTCTTCTGCCAAT 170  
 QY 111 ACTTGGAGGGATGAGTGTGTTGTTAAAGAACTGCCAATCTCTTCTGCCAAT 170  
 QY 122 GGCCTGTTCCTTCGATGCGCTCTTTCCTTGGGACATCCCTTAAGGATCTTGA 181  
 DB GGCCTGTTCCTTCGATGCGCTCTTTCCTTGGGACATCCCTTAAGGATCTTGA 230  
 QY 171 GGCCTGTTCCTTCGATGCGCTCTTTCCTTGGGACATCCCTTAAGGATCTTGA 230  
 QY 182 CATTAATTAATAATGTTATTTGATGAAATTTGATGAGTGAAGAGATGAGG 241  
 DB CATTAATTAATAATGTTATTTGATGAAATTTGATGAGTGAAGAGATGAGG 290  
 QY 231 CATTAATTAATAATGTTATTTGATGAAATTTGATGAGTGAAGAGATGAGG 290  
 QY 242 TCAAAATGAG 301  
 DB TCAAAATGAG 350  
 QY 291 TCAAAATGAG 350  
 DB TCAAAATGAG 410  
 QY 302 GTCAATGGAATACCATTAAGTTAATTAATTAATTAATTAATTAATTAAT 361  
 DB GTCAATGGAATACCATTAAGTTAATTAATTAATTAATTAATTAATTAAT 410  
 QY 351 GTCAATGGAATACCATTAAGTTAATTAATTAATTAATTAATTAATTAAT 410  
 DB GTCAATGGAATACCATTAAGTTAATTAATTAATTAATTAATTAATTAAT 421  
 QY 362 TCAAGCTGTGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 421  
 DB TCAAGCTGTGCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 470  
 QY 411 TCAAGCTGTGCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 470  
 DB TCAAGCTGTGCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 481  
 QY 422 AGAACTTGGCTTCACTTAATTAATTAATTAATTAATTAATTAATTAAT 481  
 DB AGAACTTGGCTTCACTTAATTAATTAATTAATTAATTAATTAATTAAT 530  
 QY 471 AGAACTTGGCTTCACTTAATTAATTAATTAATTAATTAATTAATTAAT 530  
 DB AGAACTTGGCTTCACTTAATTAATTAATTAATTAATTAATTAATTAAT 541  
 QY 482 AGAACTTGGCTTCACTTAATTAATTAATTAATTAATTAATTAATTAAT 541  
 DB AGAACTTGGCTTCACTTAATTAATTAATTAATTAATTAATTAATTAAT 590  
 QY 531 AGAACTTGGCTTCACTTAATTAATTAATTAATTAATTAATTAATTAAT 590  
 DB AGAACTTGGCTTCACTTAATTAATTAATTAATTAATTAATTAATTAAT 601  
 QY 542 GGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 601  
 DB GGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 650

QY 602 ATGACAGATGACCTGGAAGGCTAGAGACCTGAGTTCAAGTTACCTGACACAAAGAAAT 661  
 DB ATGACAGATGACCTGGAAGGCTAGAGACCTGAGTTCAAGTTACCTGACACAAAGAAAT 710  
 QY 662 CAGTTTCTGATTCATATGATGACAGAGGCTGACAGGACATGAGAAATGATCATC 721  
 DB CAGTTTCTGATTCATATGATGACAGAGGCTGACAGGACATGAGAAATGATCATC 770  
 QY 711 CAGTTTCTGATTCATATGATGACAGAGGCTGACAGGACATGAGAAATGATCATC 770  
 DB CAGTTTCTGATTCATATGATGACAGAGGCTGACAGGACATGAGAAATGATCATC 781  
 QY 722 TGCCTTAAGCTCTTGTGATGACATCTGTTTTCATGACCTGATTAATGCTTT 781  
 DB TGCCTTAAGCTCTTGTGATGACATCTGTTTTCATGACCTGATTAATGCTTT 830  
 QY 771 TGCCTTAAGCTCTTGTGATGACATCTGTTTTCATGACCTGATTAATGCTTT 830  
 DB TGCCTTAAGCTCTTGTGATGACATCTGTTTTCATGACCTGATTAATGCTTT 841  
 QY 782 TCCCTGTTTAAGAAACCTTATTTTCTGAAACCTCTGATGATGATGATGATGAT 841  
 DB TCCCTGTTTAAGAAACCTTATTTTCTGAAACCTCTGATGATGATGATGATGAT 890  
 QY 831 TCCCTGTTTAAGAAACCTTATTTTCTGAAACCTCTGATGATGATGATGATGAT 890  
 DB TCCCTGTTTAAGAAACCTTATTTTCTGAAACCTCTGATGATGATGATGATGAT 901  
 QY 842 CATGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901  
 DB CATGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 950  
 QY 891 CATGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 950  
 DB CATGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961  
 QY 902 GAATCTTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961  
 DB GAATCTTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1010  
 QY 951 GAATCTTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1010  
 DB GAATCTTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021  
 QY 962 GTGCTTGTGAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGAT 1021  
 DB GTGCTTGTGAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGAT 1070  
 QY 1011 GTGCTTGTGAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGAT 1070  
 DB GTGCTTGTGAAAGTACATGATGATGATGATGATGATGATGATGATGATGATGAT 1081  
 QY 1022 CTGCTTACAGCCCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081  
 DB CTGCTTACAGCCCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1130  
 QY 1071 CTGCTTACAGCCCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1130  
 DB CTGCTTACAGCCCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1141  
 QY 1082 AATTTTAAATTTCTTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 1141  
 DB AATTTTAAATTTCTTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 1190  
 QY 1131 AATTTTAAATTTCTTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 1190  
 DB AATTTTAAATTTCTTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 1201  
 QY 1142 GCTTAAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1201  
 DB GCTTAAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1250  
 QY 1191 GCTTAAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1250  
 DB GCTTAAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1261  
 QY 1202 CTTTCCCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1261  
 DB CTTTCCCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1310  
 QY 1251 CTTTCCCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1310  
 DB CTTTCCCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1321  
 QY 1262 CTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1321  
 DB CTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1370  
 QY 1311 CTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1370  
 DB CTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1381  
 QY 1322 ACATTTGCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1381  
 DB ACATTTGCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1430  
 QY 1371 ACATTTGCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1430  
 DB ACATTTGCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1441  
 QY 1382 AAAGAGCTCAG 1441  
 DB AAAGAGCTCAG 1490  
 QY 1431 AAAGAGCTCAG 1490  
 DB AAAGAGCTCAG 1501  
 QY 1442 GTTGTGAGATTTCTGAGCAATTAATGATGATGATGATGATGATGATGATGAT 1501  
 DB GTTGTGAGATTTCTGAGCAATTAATGATGATGATGATGATGATGATGATGAT 1550  
 QY 1491 GTTGTGAGATTTCTGAGCAATTAATGATGATGATGATGATGATGATGATGAT 1550  
 DB GTTGTGAGATTTCTGAGCAATTAATGATGATGATGATGATGATGATGATGAT 1561  
 QY 1502 AACAAAG 1561  
 DB AACAAAG 1621  
 QY 1551 AACAAAG 1621  
 DB AACAAAG 1670  
 QY 1611 GTTGAACAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1670  
 DB GTTGAACAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1681  
 QY 1622 CAG 1681  
 DB CAG 1730  
 QY 1671 CAG 1730  
 DB CAG 1741  
 QY 1682 ATAGTAG 1741

Db	1731	ATAGTACCGAGTGGCTGCACATGGCTATTCCAAATTTAAG-TGATTAAATTAAAGTAAAGA	17300
OY	1742	TAATAAATTAGCCTTCAGTACGGTTACCCAGCATGTAACTAAGGCTTACCACTCAGC	18010
Db	1791	TAATAATTAGCTTCAGTAGCGGTACCACTGTAACCTAAGAGGCTACCCATCAAGACG	18500
OY	1802	GTGCATAATATAGAAATTTCTCTTTAAACAGAAAGTTCTATTGGAAACAATGTTCTAGA	18610
Db	1851	GTGCATAATATAGAAATTTCTCTTTAAACAGAAAGTTCTATTGGAAACAATGTTCTAGA	19110
OY	1862	AAATATACACATATCTATTAATAAACAAAGACAGTAGATTTGTCTAAGCCAGGGGT	19210
Db	1911	AAATATACACATATCTATTAATAAACAAAGACAGTAGATTTGTCTAAGCCAGGGGT	19700
OY	1922	GAGGGAGATCCGATTGCCAAAGTGCTATAGAGAAAGTTTGGGGTAATAGCGTTTGTGAA	19810
Db	1971	GAGGGAGATCCGATTGCCAAAGTGCTATAGAGAAAGTTTGGGGTAATAGCGTTTGTGAA	20300
OY	1982	TCTTGATTGCGATGAAAGGCTACTCGGTGTCTAATGTGTCACTCTCACTAGACACTT	20410
Db	2031	TCTTGATTGCGATGAAAGGCTACTCGGTGTCTAATGTGTCACTCTCACTAGACACTT	20900
OY	2042	GCAATTGGGGAATTCATGTGATGTGAATTAATCTCATTAAGTAACTCTTAAGAGTCA	21010
Db	2091	GCAATTGGGGAATTCATGTGATGTGAATTAATCTCATTAAGTAACTCTTAAGAGTCA	21500
OY	2102	GTGTTTGTGGAAATTTATTTTAACTCAGTTGCATTACTTAATAGAAATTTTTCGA	21610
Db	2151	GTGTTTGTGGAAATTTATTTTAACTCAGTTGCATTACTTAATAGAAATTTTTCGA	2210
OY	2162	ATACATTAACATGTTATTCAATCCATTAAGTGCAATATTTTGTACGCTCTGAAAAACA	22210
Db	2211	ATACATTAACATGTTATTCAATCCATTAAGTGCAATATTTTGTACGCTCTGAAAAACA	22700
OY	2222	GAAATGAATTAGATATGCAAGCGCTCCCTCAAGCTGTTCACATCCAGATACAGAGACA	22810
Db	2271	GAAATGAATTAGATATGCAAGCGCTCCCTCAAGCTGTTCACATCCAGATACAGAGACA	23300
OY	2282	GTCTATTCCAAAATAGCTAGACTCGAGAGAAAGTTATAGTGCACTTACCAAAAAAG	23410
Db	2331	GTCTATTCCAAAATAGCTAGACTCGAGAGAAAGTTATAGTGCACTTACCAAAAAAG	23900
OY	2342	TGCAGATATATTAATGTACGACAGTAAAGTGGGAGAGTTTCTTTATATGTGAAAAAG	24010
Db	2391	TGCAGATATATTAATGTACGACAGTAAAGTGGGAGAGTTTCTTTATATGTGAAAAAG	24500
OY	2402	AGGGAATATTTTGTGCTTTGAGGATAGACAGATGTGAATATGGCGAGATGGAGTTT	24610
Db	2451	AGGGAATATTTTGTGCTTTGAGGATAGACAGATGTGAATATGGCGAGATGGAGTTT	25100
OY	2462	AAAACATTTCTGTGTGAGGCGAGATATGATCCAGGCGACAGAGCACAGAAAAATAT	25210
Db	2511	AAAACATTTCTGTGTGAGGCGAGATATGATCCAGGCGACAGAGCACAGAAAAATAT	25700
OY	2522	GCAACTTAGAGAAAGTGCATGAAGGGGACGAGTTGTAATATTTCACTAAATGTAG	25810
Db	2571	GCAACTTAGAGAAAGTGCATGAAGGGGACGAGTTGTAATATTTCACTAAATGTAG	26300
OY	2582	TGAGAAAGATTTGTATCAACAACCTGAAGTTTGGCGAGTGCATTTTGGCTCTCAG	26410
Db	2631	TGAGAAAGATTTGTATCAACAACCTGAAGTTTGGCGAGTGCATTTTGGCTCTCAG	26900
OY	2642	GAGTCAAGAAACAAGTGTCCCTTTCTCTCAAGTTATGCTCAGTGGTCCAAAGTCCAA	27010
Db	2691	GAGTCAAGAAACAAGTGTCCCTTTCTCTCAAGTTATGCTCAGTGGTCCAAAGTCCAA	27500
OY	2702	ACAACCTTCTCTCTTAAGTACTTCTCTCCCTCCATPACAATATTAAGCTTCAACA	27610
Db	2751	ACAACCTTCTCTCTTAAGTACTTCTCTCCCTCCATPACAATATTAAGCTTCAACA	28100
OY	2762	ACAATCATTAACAGGCAAGGTCAATGTGCAAAAAGCAATGCTTTTCTTAAGCTTATG	28210

Dh	2811	ACATCATTTAAACAGCAGCTCATGCTGACAAAGGCAATGCTTTCTCTACACTTCATG	2810
Oy	2822	TACGTTATTATATACAAATTTCTGCTTAAAGACTCTAAAGTCTTGAAAAGTTTCCAC	2881
Dh	2871	TACGTTATTATATATACAAATTTCTGCTTAAAGACTCTAAAGTCTTGAAAAGTTTCCAC	2930
Oy	2882	TTGCACATCAAAAGATATATATCATGCAATTTGTAATGAACTTAGTCCCTTAAGAAATA	2941
Dh	2931	TTGCACATCAAAAGATATATATCATGCAATTTGTAATGAACTTAGTCCCTTAAGAAATA	2990
Oy	2942	AGGATTAATCTAAATATATAGAACTTAATATGTAATATATATATGTAATGCACTTATTT	3001
Dh	2991	AGGATTAATCTAAATATATAGAACTTAATATGTAATATATATATGTAATGCACTTATTT	3050
Oy	3002	TTCACTTGATCGTGTATGTTGATCGTACTGCTGTTCTGTTGAATCTTAGAGACTTTGC	3061
Dh	3051	TTCACTTGATCGTGTATGTTGATCGTACTGCTGTTCTGTTGAATCTTAGAGACTTTGC	3110
Oy	3062	CTCTTTTTCCTGGGTCACCTCGCCACTTAATTTTCCATATGAAATAGAGGCAATCTTT	3121
Dh	3111	CTCTTTTTCCTGGGTCACCTCGCCACTTAATTTTCCATATGAAATAGAGGCAATCTTT	3170
Oy	3122	TTCATTAATCTTAATTTAAATTTGTTGCCATTAATTTCTGTTCTCTTAGCTTAGTA	3181
Dh	3171	TTCATTAATCTTAATTTAAATTTGTTGCCATTAATTTCTGTTCTCTTAGCTTAGTA	3230
Oy	3182	CTTTAGGATTTTAAATATTAACAATATTTGAATCATGCAATACGTTAAATGATATATTT	3241
Dh	3231	CTTTAGGATTTTAAATATTAACAATATTTGAATCATGCAATACGTTAAATGATATATTT	3280
Oy	3242	AAATAGTAGGCTTAATTAACCTTTTAAATTTTAAATAATATGATGATGATGATGCTC	3301
Dh	3291	AAATAGTAGGCTTAATTAACCTTTTAAATTTTAAATAATATGATGATGATGATGCTC	3350
Oy	3302	ATGCTCTTAATCTCCACACTTTGGGAAAGCCGGTCCGGAAGATAGCTTAGTCCAGCAGT	3361
Dh	3351	ATGCTCTTAATCTCCACACTTTGGGAAAGCCGGTCCGGAAGATAGCTTAGTCCAGCAGT	3410
Oy	3362	TTGAGACCGTAGGGCAACAAGCAAGACCCCATATCTTAAATAAACCAAAACAAACAA	3421
Dh	3411	TTGAGACCGTAGGGCAACAAGCAAGACCCCATATCTTAAATAAACCAAAACAAACAA	3470
Oy	3422	ATTACCTGGGTAATGTTGTGCTACCTGTATGTCGAAGCTACACAGAGAGCTTAGAGCAAA	3481
Dh	3471	ATTACCTGGGTAATGTTGTGCTACCTGTATGTCGAAGCTACACAGAGAGCTTAGAGCAAA	3530
Oy	3482	GGATCACTTAAGCCCTAGAGAGTTGAGCTGCAATGATCCATTAACGCGCTGCTACATCA	3541
Dh	3531	GGATCACTTAAGCCCTAGAGAGTTGAGCTGCAATGATCCATTAACGCGCTGCTACATCA	3590
Oy	3542	GTCTGGGTGACAGTGCAAGAGCTGTCTCAAAATATATATATATATATATATATATCTTTA	3601
Dh	3591	GTCTGGGTGACAGTGCAAGAGCTGTCTCAAAATATATATATATATATATATATATCTTTA	3650
Oy	3602	AAAAACAATAATTAATTAATTTTAAATAACAAACAATAGATATGTTTGCAATATGAT	3661
Dh	3651	AAAAACAATAATTAATTAATTTTAAATAACAAACAATAGATATGTTTGCAATATGAT	3710
Oy	3662	TATTTGGAGTCTATATCTCTGGAATTAATTTAAATAATTTAGAAAGTTCTTCTCAT	3721
Dh	3711	TATTTGGAGTCTATATCTCTGGAATTAATTTAAATAATTTAGAAAGTTCTTCTCAT	3770
Oy	3722	TTCTCTAGAGAGTGGAATATGTAATATCAGAGCTAGAGAGAACTTAGAGGCTCCGCACTC	3781
Dh	3771	TTCTCTAGAGAGTGGAATATGTAATATCAGAGCTAGAGAGAACTTAGAGGCTCCGCACTC	3830
Oy	3782	CAAAATGTGTGTCAAAGACCAAGAGCACTAAATGTAACCTTGGAAGCTGTTAAATATGCA	3841
Dh	3831	CAAAATGTGTGTCAAAGACCAAGAGCACTAAATGTAACCTTGGAAGCTGTTAAATATGCA	3890
Oy	3842	GTCTTAGGCTCTACCCCAAGACTATAGAACCAAGAACTTGCAATTAACAAGATTTCTAGGTG	3901
Dh	3891	GTCTTAGGCTCTACCCCAAGACTATAGAACCAAGAACTTGCAATTAACAAGATTTCTAGGTG	3950





XX 29-JUN-1999; 99JP-00183195.

PA (NIPP) NIPPON KAYAKU KK.  
PA (SAKA/) SAKAI T.

PA (SAKA/) SAKAI T.

PI Sakai T, Kagaya S, Sato T, Sukenaga Y, Fujii H,  
XX

DR WPI; 2001-112452/12.  
XY

PT New DNA sequences encoding the human p51 promoter domain for screening  
PT for compounds that modify p51 promoter activity and for diagnosing and  
PT treating cancer.

25 Claim 1(1); Page 31-34; 60pp; Japanese  
xx

This invention describes novel DNA sequences (1) encoding the human p51 promoter domain (which may include the 5'-untranslated sequence) or derived from it by addition, deletion and/or substitution of one or more bases. The invention also describes (1) expression plasmids including (1) (1) host cells transformed by (1); (3) DNA probes binding to all or part of (1); (4) cloning (1) using (1); (5) DNA and RNA corresponding to (1) inhibiting the expression of p51, which includes the antisense sequence of (1); (6) screening compounds for their ability to modify p51 promoter activity, by observing their effect on cells transformed by (1); (7) compounds identified by (6); and (8) drug compositions containing (7). The products of the invention have cytostatic activity and can be used for gene therapy. (1) is used to screen for compounds that modify p51 promoter activity. (1) and RNA corresponding to it are used to inhibit the expression of p51. (1) and the identified compounds are used for the diagnosis and treatment of cancer.

Sequence 5676 BP; 1784 A; 1053 C; 1116 G; 1723 T; 0 U; 0 Other;

Query Match	95.2%	Score 5676	DB 5	Length 5676
Best Local Similarity	100.0%	Pred. No. 0		
Matches 5676	0	Mismatches 0	Indels 0	Gaps 0

Qy	1	AAGCTGTGAAAGGATGCTCGAAGAAAGAAAGCCACCCGACATTGGCTTGTGGACATGGGGT	60
Db	1	CAGCGTTCAAGGAGTCTCTGGAAAAGAAAGCCACCCGACATTGGCTTGTGGACATGGGGT	60
Qy	61	GACTTGGAGGATACGAGTTTGTCTGTAAAGAACTGGCACTCTTCTGCGCCCAAT	120
Db	61	GACTTGGAGGATACGAGTTTGTCTGTAAAGAACTGGCACTCTTCTGCGCCCAAT	120
Qy	121	TGGCTCTGTTCCTTGCATGCCCTCTTCTCTTGGACATCCGCTTAAAGCATCTTCTG	180
Db	121	TGGCTCTGTTCCTTGCATGCCCTCTTCTCTTGGACATCCGCTTAAAGCATCTTCTG	180
Qy	181	ACATTAACTTAATTAATGTTTATTTGATGATTTCACTGACCGGAAAGAGATGGAG	240
Db	181	ACATTAACTTAATTAATGTTTATTTGATGATTTCACTGACCGGAAAGAGATGGAG	240
Qy	241	GTCAAATCAGAAAGACAGATGGCTTAAGCTTGGAAATGCATCTTCTTTTCATGAATTA	300
Db	241	GTCAAATCAGAAAGACAGATGGCTTAAGCTTGGAAATGCATCTTCTTTTCATGAATTA	300
Qy	301	AGTCAATTGGAATACCAATCACTTAAGTTAGTTAGTACGCCAGACTTACCTCTAATGAT	360
Db	301	AGTCAATTGGAATACCAATCACTTAAGTTAGTTAGTACGCCAGACTTACCTCTAATGAT	360
Qy	361	GTCAGACTGTAGCAAAATATTAGAGTCCMAAGTTGGAAAGATGTAGACAGATCTCTCCATGA	420
Db	361	GTCAGACTGTAGCAAAATATTAGAGTCCMAAGTTGGAAAGATGTAGACAGATCTCTCCATGA	420
Qy	421	CAGAACTTGTGGCTTCCACTTTTCTAAATTAAGATTGTGTGGTTAGACGTCGACGCAATGA	480
Db	421	CAGAACTTGTGGCTTCCACTTTTCTAAATTAAGATTGTGTGGTTAGACGTCGACGCAATGA	480
Qy	481	CAGAAAGCTGTCAATTAATTAATTAATTCGCAATCACTCACTTTCATTTGAGCAATGT	540
Db	481	CAGAAAGCTGTCAATTAATTAATTAATTCGCAATCACTCACTTTCATTTGAGCAATGT	540

OY	541	TGGTAAATTATTCACCAACACCTCCCTCTCTGAAAGGCTCTTCTCGCTCCCTCCACTATA	600
Db	541	TGGTAAATTATTCACCAACACCTCCCTCTCTGAAAGGCTCTTCTCGCTCCCTCCACTATA	600
OY	601	ATATGAGATAGACCTGGAAAGGCTTAGAACCTGAGGCTGACGTTAACCTTCACACAAAGAT	660
Db	601	ATATGAGATAGACCTGGAAAGGCTTAGAACCTGAGGCTGACGTTAACCTTCACACAAAGAT	660
OY	661	TCAGTTTCTGTATCTCATAGTCAAGAGCTGACAGAGCTTACGAAACATGCAAGATCAT	720
Db	661	TCAGTTTCTGTATCTCATAGTCAAGAGCTGACAGAGCTTACGAAACATGCAAGATCAT	720
OY	721	CTGCTTAAAGCCTTGTGTGGGATCTGTGTGTTTCACTGCGCCTGACCTATATGCTCT	780
Db	721	CTGCTTAAAGCCTTGTGTGGGATCTGTGTGTTTCACTGCGCCTGACCTATATGCTCT	780
OY	781	TTCTGTGGTTAACAGAACTTTATTTTCTGTAAAACTCTGTGTCAGTCAAGTAAAGG	840
Db	781	TTCTGTGGTTAACAGAACTTTATTTTCTGTAAAACTCTGTGTCAGTCAAGTAAAGG	840
OY	841	CCATCACTCCAGATGATTCAGGCTCTCTGTGCGCAACATGSCATCTTTCTTTGGAAAT	900
Db	841	CCATCACTCCAGATGATTCAGGCTCTCTGTGCGCAACATGSCATCTTTCTTTGGAAAT	900
OY	901	TGAACTTAAAGCTGAATACCTGAAGTTCAAAAAAGCTTTGAACTTGACTTACAGCCTAC	960
Db	901	TGAACTTAAAGCTGAATACCTGAAGTTCAAAAAAGCTTTGAACTTGACTTACAGCCTAC	960
OY	961	AGTGGCTTTCCAAAGTGAATGTGCATTCCTATTTCTTAAAGTCCCTGAATTTATATTAAT	1020
Db	961	AGTGGCTTTCCAAAGTGAATGTGCATTCCTATTTCTTAAAGTCCCTGAATTTATATTAAT	1020
OY	1021	CCGTGGTTACAGCCCTTCTGAGATGTGTGTTTTTTTCCAACTGTCTCTATATAGCTGT	1080
Db	1021	CCGTGGTTACAGCCCTTCTGAGATGTGTGTTTTTTTCCAACTGTCTCTATATAGCTGT	1080
OY	1081	GAAATTTCAATTTCTTTTCAATACATTTTCAGTGTGTTGTTGTGTGTGTGTGT	1140
Db	1081	GAAATTTCAATTTCTTTTCAATACATTTTCAGTGTGTTGTTGTGTGTGTGTGTGT	1140
OY	1141	GACTTTAGTAGGAGCAATCAGTTCTGTGTTTATACCAAGAAATCCGATTTGATACA	1200
Db	1141	GACTTTAGTAGGAGCAATCAGTTCTGTGTTTATACCAAGAAATCCGATTTGATACA	1200
OY	1201	TCTCTCCCTTTAAAAATTAAGTATCTAAGGCTCAAGAGAGTAAAGCTCACTCGAGAG	1260
Db	1201	TCTCTCCCTTTAAAAATTAAGTATCTAAGGCTCAAGAGAGTAAAGCTCACTCGAGAG	1260
OY	1261	TCTGGAGTAAAGTGTGATACAGAGCTGTATACCCCAAGGTAAAGCACTGTATTACAC	1320
Db	1261	TCTGGAGTAAAGTGTGATACAGAGCTGTATACCCCAAGGTAAAGCACTGTATTACAC	1320
OY	1321	AACATTTGCTCTCTCTTCAGAGTTATAGCAGCTTGTGAAGAAAGAGTACTATTTTGC	1380
Db	1321	AACATTTGCTCTCTCTTCAGAGTTATAGCAGCTTGTGAAGAAAGAGTACTATTTTGC	1380
OY	1381	CAAGAGCTTCAGAGGACCAAGAAACAATTTCTGGGATATGTGATATTGAACCTTAAAA	1440
Db	1381	CAAGAGCTTCAGAGGACCAAGAAACAATTTCTGGGATATGTGATATTGAACCTTAAAA	1440
OY	1441	AGTTGTGTGACCTTGTGCACTAATTTGTGTATCTTAAAGCAAGATTTCACTTTAAATAGT	1500
Db	1441	AGTTGTGTGACCTTGTGCACTAATTTGTGTATCTTAAAGCAAGATTTCACTTTAAATAGT	1500
OY	1501	AAAACAAACAAACAGAGATCCAGAGGTTCCAGGCTATATATATATATATATATATATAT	1560
Db	1501	AAAACAAACAAACAGAGATCCAGAGGTTCCAGGCTATATATATATATATATATATATATAT	1560
OY	1561	AGTTGACACATGATATATATCTCAAGAACATGTGCATATATACAAAAAAATTCAGAC	1620
Db	1561	AGTTGACACATGATATATATCTCAAGAACATGTGCATATATACAAAAAAATTCAGAC	1620

[illegible]

Dh	2701	AAACGCTTTCCTCTTAAGATCACTTCTCTCCCTCCATACAAATCTAAAGCTTCACAA	2760
Qy	2761	AACTCATTTTAAACAGGCAAGTCATGCTCAGAAAGGCATTTGCTTTCATAGCTCTAT	2820
Dh	2761	AACTCATTTTAAACAGGCAAGTCATGCTCAGAAAGGCATTTGCTTTCATAGCTCTAT	2820
Qy	2821	GTACGTTATTTATTTACATTTTCGCTTAAAGACTCTAAAGCTTTGAAAAGTTTCAC	2880
Dh	2821	GTACGTTATTTATTTACATTTTCGCTTAAAGACTCTAAAGCTTTGAAAAGTTTCAC	2880
Qy	2881	CTTGCAATCAAAAGATTAATTCATGCAATTTGATATAGTAACTAGTCCCTAAAGAAAT	2940
Dh	2881	CTTGCAATCAAAAGATTAATTCATGCAATTTGATATAGTAACTAGTCCCTAAAGAAAT	2940
Qy	2941	AAGATGAACTATAAATTTAAGAGATTAATGATATTAATTAATTAATGATGCACTATT	3000
Dh	2941	AAGATGAACTATAAATTTAAGAGATTAATGATATTAATTAATTAATGATGCACTATT	3000
Qy	3001	TTTCACCTGATGATAGTATGCTTGATGCTATGCTGCTCTGTTGTAATTTAGAGAGTTTG	3060
Dh	3001	TTTCACCTGATGATAGTATGCTGATGCTATGCTGCTCTGTTGTAATTTAGAGAGTTTG	3060
Qy	3061	CTCTTTTTCGCGGTGAACCTCTGCGCAATTTATTCATATGCAATTAAGAGCAATCTT	3120
Dh	3061	CTCTTTTTCGCGGTGAACCTCTGCGCAATTTATTCATATGCAATTAAGAGCAATCTT	3120
Qy	3121	TTTCATTAATTAATTTTAAATTTGTCGCAATTTAATTTCTGTTCTCTTATGCTTAGTA	3180
Dh	3121	TTTCATTAATTAATTTTAAATTTGTCGCAATTTAATTTCTGTTCTCTTATGCTTAGTA	3180
Qy	3181	ACCTTGGAGTTTAAATTAACATGATTTGAAATCATGACATAGCTTTAAATGAATATT	3240
Dh	3181	ACCTTGGAGTTTAAATTAACATGATTTGAAATCATGACATAGCTTTAAATGAATATT	3240
Qy	3241	TAAATAGCTAGCTTTAAACCTTTTAAATTTTAAATTTTAAATTAATAGATGATGCTGCT	3300
Dh	3241	TAAATAGCTAGCTTTAAACCTTTTAAATTTTAAATTTTAAATTAATAGATGATGCTGCT	3300
Qy	3301	CATGCTGTAAATCCCAACCTTTGGGAGCCGGTCTGGAGATTAAGCTTAGTCAAGCAG	3360
Dh	3301	CATGCTGTAAATCCCAACCTTTGGGAGCCGGTCTGGAGATTAAGCTTAGTCAAGCAG	3360
Qy	3361	TTTGAACACGATCGAGGSCAACACAGCAAGCCCATATCTTAAAAAAAACAAAACAA	3420
Dh	3361	TTTGAACACGATCGAGGSCAACACAGCAAGCCCATATCTTAAAAAAAACAAAACAA	3420
Qy	3421	AATTAAGCTGGATAGTGTGCTCACTCTATGTCACAGCTACACAGGAAGCTGAGCGACA	3480
Dh	3421	AATTAAGCTGGATAGTGTGCTCACTCTATGTCACAGCTACACAGGAAGCTGAGCGACA	3480
Qy	3481	AGATCACTTGAACCCAGAGAGGTTGAGGCTGACATGATCATGAAACGGCTGCTACATC	3540
Dh	3481	AGATCACTTGAACCCAGAGAGGTTGAGGCTGACATGATCATGAAACGGCTGCTACATC	3540
Qy	3541	AGTCTGGGTGACAGTCCAAAGAGCTGCTCAAAAATTAATTAATTAATTAATTAATCTTT	3600
Dh	3541	AGTCTGGGTGACAGTCCAAAGAGCTGCTCAAAAATTAATTAATTAATTAATTAATCTTT	3600
Qy	3601	AAAAAACAATAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA	3660
Dh	3601	AAAAAACAATAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA	3660
Qy	3661	TTATTTGGGAGTCTATATCCCTGGAAGTTAATTTAAATTTTGAAGAGTCTCTCTCA	3720
Dh	3661	TTATTTGGGAGTCTATATCCCTGGAAGTTAATTTAAATTTTGAAGAGTCTCTCTCA	3720
Qy	3721	TTCTCTAGAGAGCTGCAATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT	3780
Dh	3721	TTCTCTAGAGAGCTGCAATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT	3780
Qy	3781	CGAAGTGTGTCAGAGCAAGCAAGCATGAAGTAATCTGGAAAGCTGTATGAATTAAGAG	3840
Dh	3781	CGAAGTGTGTCAGAGCAAGCAAGCATGAAGTAATCTGGAAAGCTGTATGAATTAAGAG	3840

Db 3781 CCAAGTGTGTCCAGACACAGACATCAAGTACCTGGAAAGTGTAGAAATGCAG 3840  
 QY 3841 AGTTTATGGCTTCAACCCAGACCTTCTGACCAAGATCTGTATTAACAGATTTCTAGT 3900  
 Db 3841 AGTTTATGGCTTCAACCCAGACCTTCTGACCAAGATCTGTATTAACAGATTTCTAGT 3900  
 QY 3901 GCTCAGCGGACATTTAAACTTGTAGAGCTCTGACATAGAAATCTCACTCCACTTTC 3960  
 Db 3901 GCTCAGCGGACATTTAAACTTGTAGAGCTCTGACATAGAAATCTCACTCCACTTTC 3960  
 QY 3961 ATTATTAATGSAATCTTGGGCTGTGTGACAGAGAAATGATTAATTTTAAATTTTACAA 4020  
 Db 3961 ATTATTAATGSAATCTTGGGCTGTGTGACAGAGAAATGATTAATTTTAAATTTTACAA 4020  
 QY 4021 CCTCTATTTAGTCACTATATTTGCTATAGCAAGGAGAAAGCCAACTCTTTAACT 4080  
 Db 4021 CCTCTATTTAGTCACTATATTTGCTATAGCAAGGAGAAAGCCAACTCTTTAACT 4080  
 QY 4081 GCATTTAACAAATCTATTAATTAATTAAGTAAATCTTCCCTTAAGTTTACATTTTG 4140  
 Db 4081 GCATTTAACAAATCTATTAATTAATTAAGTAAATCTTCCCTTAAGTTTACATTTTG 4140  
 QY 4141 TGGAGCAGCTGTATTAATTTGGTGGGCTGAGGCGGCTGTGTGATTTTCAATTT 4200  
 Db 4141 TGGAGCAGCTGTATTAATTTGGTGGGCTGAGGCGGCTGTGTGATTTTCAATTT 4200  
 QY 4201 CACGATGTTAGCGCTCTGGGCTTAAAGTAAAGAGAAATGTCAGTTTAAATAGCT 4260  
 Db 4201 CACGATGTTAGCGCTCTGGGCTTAAAGTAAAGAGAAATGTCAGTTTAAATAGCT 4260  
 QY 4261 TCTCCCTTCATCTGGCTGAGACAAATTAATTTTAAATTTTAAATTTTAAATTTTAAAT 4320  
 Db 4261 TCTCCCTTCATCTGGCTGAGACAAATTAATTTTAAATTTTAAATTTTAAATTTTAAAT 4320  
 QY 4321 TAGATTTACTTAAGGAAATGTCAAATTTCTGAAAGGCTTTAGATTTCTCAAC 4380  
 Db 4321 TAGATTTACTTAAGGAAATGTCAAATTTCTGAAAGGCTTTAGATTTCTCAAC 4380  
 QY 4381 TTTGACATCTAGATGTCACCTATTTTACAGTGTGTCTGTGATAGGAGGAGTAAAGGGA 4440  
 Db 4381 TTTGACATCTAGATGTCACCTATTTTACAGTGTGTCTGTGATAGGAGGAGTAAAGGGA 4440  
 QY 4441 AGATGTAATCTCAACAGTGTAGGACCTTGTAGTACAGAGTGTGTGTGTGTGTGTGTGTGT 4500  
 Db 4441 AGATGTAATCTCAACAGTGTAGGACCTTGTAGTACAGAGTGTGTGTGTGTGTGTGTGTGT 4500  
 QY 4501 TTTGAGTCTATCTTAAGT 4560  
 Db 4501 TTTGAGTCTATCTTAAGT 4560  
 QY 4561 ACCAGATTAAGT 4620  
 Db 4561 ACCAGATTAAGT 4620  
 QY 4621 TCTGATTAATCTCAATTTGGGATATATTAAGATTTTATTTTATTTTATTTTATTTTATTT 4680  
 Db 4621 TCTGATTAATCTCAATTTGGGATATATTAAGATTTTATTTTATTTTATTTTATTTTATTT 4680  
 QY 4681 GTCAAGTCAAGGAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAA 4740  
 Db 4681 GTCAAGTCAAGGAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAA 4740  
 QY 4741 CCGACTGAGAGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4800  
 Db 4741 CCGACTGAGAGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4800  
 QY 4801 GTGACTGTGAGATTTTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4860  
 Db 4801 GTGACTGTGAGATTTTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4860  
 QY 4861 TTTACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4920  
 Db 4861 TTTACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4920

QY 4921 AGTTTATTTTCTTTTATTTGCGGTCAAGCAAACTTTCTAAGGAGATGTGAAGGATATCT 4980  
 Db 4921 AGTTTATTTTCTTTTATTTGCGGTCAAGCAAACTTTCTAAGGAGATGTGAAGGATATCT 4980  
 QY 4981 CTTTCTCTAGCTGAGAGAGAGT 5040  
 Db 4981 CTTTCTCTAGCTGAGAGAGAGT 5040  
 QY 5041 GTCTTTGCTATTTAGATTTGTACCAACAGGCGGTTGCTGAAAGGAACTGAAGG 5100  
 Db 5041 GTCTTTGCTATTTAGATTTGTACCAACAGGCGGTTGCTGAAAGGAACTGAAGG 5100  
 QY 5101 CCGGAGAGGAGGAAATAGATGAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAAC 5160  
 Db 5101 CCGGAGAGGAGGAAATAGATGAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAAC 5160  
 QY 5161 ACAAACATTTTGGCCCGCAAAATGTCAAGAAATCTCAAAATCAAAACGATTCGAG 5220  
 Db 5161 ACAAACATTTTGGCCCGCAAAATGTCAAGAAATCTCAAAATCAAAACGATTCGAG 5220  
 QY 5221 TACAAGAGTGTATGTATGTGAGCTGAGCAAGGCTGACACTCATGAGCTCAGTTGATCA 5280  
 Db 5221 TACAAGAGTGTATGTATGTGAGCTGAGCAAGGCTGACACTCATGAGCTCAGTTGATCA 5280  
 QY 5281 AAGTCCAGGCTGCAAAATTAATCTGATGCCATTCATATGCCAGATCCAAATCAGACA 5340  
 Db 5281 AAGTCCAGGCTGCAAAATTAATCTGATGCCATTCATATGCCAGATCCAAATCAGACA 5340  
 QY 5341 CAGATCAAGATTTCAAGATGTGCTGAGCTGCAATTTGCCAACAAGTGTGGTACTA 5400  
 Db 5341 CAGATCAAGATTTCAAGATGTGCTGAGCTGCAATTTGCCAACAAGTGTGGTACTA 5400  
 QY 5401 TACGTCAGAGACTCTGAAGCGGTGAGAGGAGGAGAAACAAGTGAAGATGCCA 5460  
 Db 5401 TACGTCAGAGACTCTGAAGCGGTGAGAGGAGGAGAAACAAGTGAAGATGCCA 5460  
 QY 5461 GCTGTGAAGATGAGAGTGTATTAAGTGTATTAAGTGTATTAAGTGTATTAAGTGTATTAAG 5520  
 Db 5461 GCTGTGAAGATGAGAGTGTATTAAGTGTATTAAGTGTATTAAGTGTATTAAGTGTATTAAG 5520  
 QY 5521 TCAAGAAAGCTCCGCTCTTTGCCAAATTTGATTAAGAGAGAGATGCTTAACTTCTA 5580  
 Db 5521 TCAAGAAAGCTCCGCTCTTTGCCAAATTTGATTAAGAGAGAGATGCTTAACTTCTA 5580  
 QY 5581 TGTCTGATGACATTTGACCTTATGCTTTAGCCTCCGCTTTATATCTATATATAC 5640  
 Db 5581 TGTCTGATGACATTTGACCTTATGCTTTAGCCTCCGCTTTATATCTATATATAC 5640  
 QY 5641 AGTATTTGTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 5700  
 Db 5641 AGTATTTGTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 5700

RESULT 4  
 AAF25907/C  
 ID AAF25907 standard; RNA; 5676 BP.  
 XX AAF25907;  
 DT 19-APR-2001 (first entry)  
 XX  
 DB Human tumor suppressor gene p51 promoter associated RNA SEQ ID 4.  
 XX  
 KW Tumor suppressor; p51; cell death; cell proliferation; cancer;  
 XX cytotoxic; gene therapy; screening; da.  
 OS Homo sapiens.  
 XX  
 PN W0200100818-A1.  
 PD 04-JAN-2001.  
 XX

28-JUN-2000; 2000MO-JP004261.

XX 29-JUN-1999; 99JP-00183195.

XX (NIPK) NIPPON KAYAKU KK.

PA (SARA) SAKAI T.

XX Sakai T, Kagaya S, Sato T, Sukenaga Y, Fujii H;

XX WPI; 2001-112452/12.

XX New DNA sequences encoding the human p51 promoter domain for screening  
FT for compounds that modify p51 promoter activity and for diagnosing and  
FT treating cancer.

XX Disclosure; Page 42-45; 60pp; Japanese.

XX This invention describes novel DNA sequences (1) encoding the human p51  
CC promoter domain (which may include the 5'-untranslated sequence) or  
CC derived from it by addition, deletion and/or substitution of one or more  
CC bases. The invention also describes (1) expression plasmids including (1)  
CC (2) host cells transformed by (1); (3) DNA probe binding to all or  
CC part of (1); (4) cloning (1) using (3); (5) DNA (and RNA corresponding to  
CC (1) inhibiting the expression of p51, which includes the antisense  
CC sequence of (1); (6) screening compounds for their ability to modify p51  
CC promoter activity, by observing their effect on cells transformed by (1);  
CC (7) compounds identified by (6); and (8) drug compositions containing  
CC (7). The products of the invention have cytotoxic activity and can be  
CC used for gene therapy. (1) is used to screen for compounds that modify  
CC p51 promoter activity. (1) and RNA corresponding to it are used to  
CC inhibit the expression of p51. (1) and the identified compounds are used  
CC for the diagnosis and treatment of cancer

XX Sequence 5676 BP; 1723 A; 1116 C; 1053 G; 0 T; 1784 U; 0 Other;

XX Query Match 95.2%; Score 5676; DB 5; Length 5676;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 5676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CAGCTGTCAGGAGATGTCGAGAAAGAGCCACCCACATGCTTCGACACTGGCTGT	60
DB	5676	CAGCTGTCAGGAGATGTCGAGAAAGAGCCACCCACATGCTTCGACACTGGCTGT	5617
QY	61	GACCTTGGAGGGATACAGGTTGTCTGTAAAGAACTGGCAACCTCCGCGCCGAT	120
DB	5616	GACCTTGGAGGGATACAGGTTGTCTGTAAAGAACTGGCAACCTCCGCGCCGAT	5557
QY	121	TGGCTCTGTCCTTGCATGAGCCCTCTTCTTGGACACTCCCTTAAGCATCTTCTG	180
DB	5556	TGGCTCTGTCCTTGCATGAGCCCTCTTCTTGGACACTCCCTTAAGCATCTTCTG	5497
QY	181	ACATTAACCTAACTAAATGTTATTTGATGATTCATGACCTGAGAGAGATGAG	240
DB	5496	ACATTAACCTAACTAAATGTTATTTGATGATTCATGACCTGAGAGAGATGAG	5437
QY	241	GTCATATGAGAAAGACATGAGCTAGAGTTCGATCTGCTTTTCAATTGAATTA	300
DB	5436	GTCATATGAGAAAGACATGAGCTAGAGTTCGATCTGCTTTTCAATTGAATTA	5377
QY	301	AGTCATTCGATACCACTTCACTTAAGTTCTAGGCCACGCTTACTCTTAATCAT	360
DB	5376	AGTCATTCGATACCACTTCACTTAAGTTCTAGGCCACGCTTACTCTTAATCAT	5317
QY	361	GTCAGACTGAGAAATTTAGGTCGAAAGTTGAGATAGATAGATCTCTCAATGA	420
DB	5316	GTCAGACTGAGAAATTTAGGTCGAAAGTTGAGATAGATAGATCTCTCAATGA	5257
QY	421	CAGAACTTTGGCTTCACTTAATAATAGATGTTGTTGTTGAGCTGAGATGTA	480
DB	5256	CAGAACTTTGGCTTCACTTAATAATAGATGTTGTTGTTGAGCTGAGATGTA	5197
QY	481	CAGAAAGTGTATCAATTAATAATCAAGAACTGCTCTTCAATTGAGCAATGCT	540

DB	5196	CAGAAAGTGTATCAATTAATAATAACCAAACTCAGCTCTCAATTGAGCAATGCT	5137
QY	541	TGGTGAATTTATTCACACACCTCCCTCTCTGAAAGTTCTTCTCTCTCCCTACTATA	600
DB	5136	TGGTGAATTTATTCACACACCTCCCTCTCTGAAAGTTCTTCTCTCTCCCTACTATA	5077
QY	601	AATCGAGATGACCTGAGAAAGCTTAGACCTGAGCTTCACTTACCTGACACAAAGAT	660
DB	5076	AATCGAGATGACCTGAGAAAGCTTAGACCTGAGCTTCACTTACCTGACACAAAGAT	5017
QY	661	TCAATTTCTCTATCTCATAGTACAGAGCTCCAGAGCTCTAGGAAACATGCAATCAT	720
DB	5016	TCAATTTCTCTATCTCATAGTACAGAGCTCCAGAGCTCTAGGAAACATGCAATCAT	4957
QY	721	CTGCTTAAAGCTCTTGGTGGGAAATCTGTTTTCACATGGCTGAACTTATGCTCT	780
DB	4956	CTGCTTAAAGCTCTTGGTGGGAAATCTGTTTTCACATGGCTGAACTTATGCTCT	4897
QY	781	TTCTTGGTTTACAGAACTTATTTTCTTGAAAACCTCTGCTCATGCTATGCTAGC	840
DB	4896	TTCTTGGTTTACAGAACTTATTTTCTTGAAAACCTCTGCTCATGCTATGCTAGC	4837
QY	841	CCATCAGTCACATGATCAGGCTCTCCGAGCAAACTGGACATCTTCTTTGGGAAT	900
DB	4836	CCATCAGTCACATGATCAGGCTCTCCGAGCAAACTGGACATCTTCTTTGGGAAT	4777
QY	901	TGAATCTTAAAGCTGAATTAAGTTGTAAGTTGTAAGTTGTAAGTTGTAAGTTGTA	960
DB	4776	TGAATCTTAAAGCTGAATTAAGTTGTAAGTTGTAAGTTGTAAGTTGTAAGTTGTA	4717
QY	961	AGTGGCTTTCAGAAAGTACCTGATTCATTTCTTAAGTCCCTGAAATTAATTAAT	1020
DB	4716	AGTGGCTTTCAGAAAGTACCTGATTCATTTCTTAAGTCCCTGAAATTAATTAAT	4657
QY	1021	CCGCTTAAAGCCCTTTCGAGATGAGTGTGTTTTCACATGCTCTTATAGCTCT	1080
DB	4656	CCGCTTAAAGCCCTTTCGAGATGAGTGTGTTTTCACATGCTCTTATAGCTCT	4597
QY	1081	GAAATTTTCATATTTCTTCTTCAATTTTCAATTTGTTGTTGTTGTTGTTGTTGTT	1140
DB	4596	GAAATTTTCATATTTCTTCTTCAATTTTCAATTTGTTGTTGTTGTTGTTGTTGTT	4537
QY	1141	GCTTTTGAAGGACAGAAATCAATTTCTGTTGTTTATCCAGGAATCTGATGATACA	1200
DB	4536	GCTTTTGAAGGACAGAAATCAATTTCTGTTGTTTATCCAGGAATCTGATGATACA	4477
QY	1201	TCTTTTCCCTTAAATAAATAGTATCTAAGCTGAAAGAGTGGCTACTGCTGAGG	1260
DB	4476	TCTTTTCCCTTAAATAAATAGTATCTAAGCTGAAAGAGTGGCTACTGCTGAGG	4417
QY	1261	TCTGGAGTAAAGTTAGTACAGAGCTGTAATACCCAGGTTAGCCAACTGCTTACAC	1320
DB	4416	TCTGGAGTAAAGTTAGTACAGAGCTGTAATACCCAGGTTAGCCAACTGCTTACAC	4357
QY	1321	AACATTTGCTCTCTCTTCAAGATTATAGCACTCTGGAAGAAAGATCTACTATTGG	1380
DB	4356	AACATTTGCTCTCTCTTCAAGATTATAGCACTCTGGAAGAAAGATCTACTATTGG	4297
QY	1381	CAAGACCTCAGAGACCAAGAACAGTTCTGGAAATGATGATGATTAATCTTAA	1440
DB	4296	CAAGACCTCAGAGACCAAGAACAGTTCTGGAAATGATGATGATGATTAATCTTAA	4237
QY	1441	AGTTGTTGAGCTTCTGCAATATATGATCTAAGACCAAGATTTCACTTAATAGCT	1500
DB	4236	AGTTGTTGAGCTTCTGCAATATATGATCTAAGACCAAGATTTCACTTAATAGCT	4177
QY	1501	AAAGCAACCAAGAGATTCACAGGTTTACAGAGTTATATATAGTGAATTTCTGATAC	1560
DB	4176	AAAGCAACCAAGAGATTCACAGGTTTACAGAGTTATATATAGTGAATTTCTGATAC	4117
QY	1561	AGTTGACATGATATATCTCAAGAACATGCGATCAATGACCAAAAAATCCGAGAC	1620
DB	4116	AGTTGACATGATATATCTCAAGAACATGCGATCAATGACCAAAAAATCCGAGAC	4057

QY	2701	AAACCTTTCCTTCCTTAAGTACTTCTCTCCCTCCATACAATCTAAAGTCTTCACA	2766
Db	2976	AAACCTTTCCTTCCTTAAGTACTTCTCTCCCTCCATACAATCTAAAGTCTTCACA	2917
QY	2761	AAATCATTTAAACAGCAGCTCATCGTGCAGAAAGGCAATTCCTTTCTAGACTTCAT	2820
Db	2916	AAATCATTTAAACAGCAGGAGCTCATCGTGCAGAAAGGCAATTCCTTTCTAGACTTCAT	2857
QY	2821	GAACGTATTTATTTACAAATTTTCGCTTAATAAGACTCTAAAGCTTCGAAAAGTTTCCAC	2880
Db	2856	GAACGTATTTATTTACAAATTTTCGCTTAATAAGACTCTAAAGCTTCGAAAAGTTTCCAC	2797
QY	3001	TTTCACTTGATCGTGTAATGCTTGCACTGCTAGTCTGTCCTTGTAATCTAGAGATTTCG	3060
Db	2676	TTTCACTTGATCGTGTAATGCTTGCACTGCTAGTCTGTCCTTGTAATCTAGAGATTTCG	2617
QY	3061	CCCTCTTTTCCGGGTCAACTCTCGCAATTTATTTGCATATAGCAATTAAGAGCCCAATCT	3120
Db	2616	CCCTCTTTTCCGGGTCAACTCTCGCAATTTATTTGCATATAGCAATTAAGAGCCCAATCT	2557
QY	3121	TTTCATTAATTAATTTAAATTTGTTGCAATTTAATTCGTCTCTTAGCTTAGTA	3180
Db	2556	TTTCATTAATTAATTTAAATTTGTTGCAATTTAATTCGTCTCTTAGCTTAGTA	2497
QY	3181	ACTTTAGATATTTTAAATTAACAATATGTAATAATCATAGATACGTTTAAATAGATATT	3240
Db	2496	ACTTTAGATATTTTAAATTAACAATATGTAATAATCATAGATACGTTTAAATAGATATT	2437
QY	3241	TAAATACGTAGGCTATTAACCTTTTAAATTTTAAATTAATAATAGATAGTGTGCT	3300
Db	2436	TAAATACGTAGGCTATTAACCTTTTAAATTTTAAATTAATAATAGATAGTGTGCT	2377
QY	3301	CATGCTGTAAATCCCAACACTTTTGGGAAGCTGGGTGCGAGAGATAGCTTGACCAAG	3360
Db	2376	CATGCTGTAAATCCCAACACTTTTGGGAAGCTGGGTGCGAGAGATAGCTTGACCAAG	2317
QY	3361	TTTGAAGACAGTACAGGGCAACACAGCAAGACCCCATATCTAATAAAAAACAAAAACAA	3420
Db	2316	TTTGAAGACAGTACAGGGCAACACAGCAAGACCCCATATCTAATAAAAAACAAAAACAA	2257
QY	3421	AATTAACCTGGGTAATGTTGTGCTCACTGTATGTGTCAGATACAGAGAAAGCTGAGGCGA	3480
Db	2256	AATTAACCTGGGTAATGTTGTGCTCACTGTATGTGTCAGATACAGAGAAAGCTGAGGCGA	2197
QY	3481	AGGATCACTTGAGCCCAAGAGGTTGAGGCTGCAAGTATCCATGAACCGGCTGTACTAC	3540
Db	2196	AGGATCACTTGAGCCCAAGAGGTTGAGGCTGCAAGTATCCATGAACCGGCTGTACTAC	2137
QY	3541	AGTCTGGGGAAGTGCAGAAAGCTGTCTCAAAAATATATTAATTAATAATAATACTTT	3600
Db	2136	AGTCTGGGGAAGTGCAGAAAGCTGTCTCAAAAATATATTAATTAATAATAATACTTT	2077
QY	3601	AAAAAACAAAATTAATTAATTTTAAACACACACACACATAGAGATGTTTGCAATTTGA	3660
Db	2076	AAAAAACAAAATTAATTAATTTTAAACACACACACACATAGAGATGTTTGCAATTTGA	2017
QY	3661	TTATTTGGGAGTCTAATATCCCTGGAGTTTATTTAAATATATTAAGAGTTCCTCTCA	3720
Db	2016	TTATTTGGGAGTCTAATATCCCTGGAGTTTATTTAAATATATTTAGAGAGTTCCTCTCA	1957
QY	3721	TTTCTGTAGAGACGTGCAATGTTGAATATCAAGGCTTGAAGAGAAACCTAGGCTCGCACT	3780
Db	1956	TTTCTGTAGAGACGTGCAATGTTGAATATCAAGGCTTGAAGAGAAACCTAGGCTCGCACT	1897
QY	3781	CCAAAGTGTGCTCAAGACACGACGATCAAGTAACTTGAGAACGTGTTAAGAAATCG	3840

Db 1836 CCAAAAGTGTCCAGAGCACAGACATCAAGTAACTCGGAGAGCTTTGAAATGACG 1837  
 Qy 3841 AGCTTAGAGCTCAGCCCAAGACTACTGAACAGAACTGCAATTAACAAAGTTTCTAGCT 3900  
 Db 1836 AGCTTAGAGCTCAGCCCAAGACTACTGAACAGAACTGCAATTAACAAAGTTTCTAGCT 1777  
 Qy 3901 GGCTCAGGGGACATTAACCTTAGAGAGCTGCACTAGAAATCTTCACTCCACTTTC 3960  
 Db 1776 GGCTCAGGGGACATTAACCTTAGAGAGCTGCACTAGAAATCTTCACTCCACTTTC 1717  
 Qy 3961 AATTAAATGAAATCACTTGGGCTGTGTCAAGGAAATTTGATTTTAAATTTGGA 4020  
 Db 1716 AATTAAATGAAATCACTTGGGCTGTGTCAAGGAAATTTGATTTTAAATTTGGA 1657  
 Qy 4021 CCTTCATTTAGGTATCTAATTTGCTAATAGCAGGAGAAAGCAAACTCTTAACT 4080  
 Db 1656 CCTTCATTTAGGTATCTAATTTGCTAATAGCAGGAGAAAGCAAACTCTTAACT 1597  
 Qy 4081 GGAATTAACAATCTAATTAATTAATTAAGCAATCTCCCTTAAGTTTACATTTG 4140  
 Db 1596 GGAATTAACAATCTAATTAATTAATTAAGCAATCTCCCTTAAGTTTACATTTG 1537  
 Qy 4141 TGGAGCAAGCTGTGATTTGGCTGAGGCTCAGGCGGCTGTGTTGATTTCAATTT 4200  
 Db 1536 TGGAGCAAGCTGTGATTTGGCTGAGGCTCAGGCGGCTGTGTTGATTTCAATTT 1477  
 Qy 4201 CACAGATTTAGCCGCTCTGAGGCTAGTAAAGAGAGATGTCAGTTTAAATAGCT 4260  
 Db 1476 CACAGATTTAGCCGCTCTGAGGCTAGTAAAGAGAGATGTCAGTTTAAATAGCT 1417  
 Qy 4261 TCTCCCTTCATCTGCTGCTGAGCAACAATTAATTTTATGAAAACACTTTGAGT 4320  
 Db 1416 TCTCCCTTCATCTGCTGCTGAGCAACAATTAATTTTATGAAAACACTTTGAGT 1357  
 Qy 4321 TAGATTTACTTACAGGAAATGTCAAATTTCTGAAAGGGCTTTAGTTGCTCAAC 4380  
 Db 1356 TAGATTTACTTACAGGAAATGTCAAATTTCTGAAAGGGCTTTAGTTGCTCAAC 1297  
 Qy 4381 TTTGAGATCTAGTATGTCACCTAATTTTACAGAGTGTCTGATCAGAGGGTGAAGGA 4440  
 Db 1296 TTTGAGATCTAGTATGTCACCTAATTTTACAGAGTGTCTGATCAGAGGGTGAAGGA 1237  
 Qy 4441 AGATGGAATCAACATTTAGTACCGTTTATGATACAGAGTGTCTTTTCCCTG 4500  
 Db 1236 AGATGGAATCAACATTTAGTACCGTTTATGATACAGAGTGTCTTTTCCCTG 1177  
 Qy 4501 TTGAGTCTATCTAAGTACGCTTCTGAAATCAATTTCAATTTCCAAATTCACAAA 4560  
 Db 1176 TTGAGTCTATCTAAGTACGCTTCTGAAATCAATTTCAATTTCCAAATTCACAAA 1117  
 Qy 4561 ACCAGATTAAGTTTACAGGCTATGTCAGAAAGAAATTAATTTTGTGTAGACTT 4620  
 Db 1116 ACCAGATTAAGTTTACAGGCTATGTCAGAAAGAAATTAATTTTGTGTAGACTT 1057  
 Qy 4621 TCCGTATTTAATCACTGATTTTGGAAATTAATGAAATTTTATGTTTCTTTGAGTAG 4680  
 Db 1056 TCCGTATTTAATCACTGATTTTGGAAATTAATGAAATTTTATGTTTCTTTGAGTAG 997  
 Qy 4681 GTCAAGTCAAGCAAAACCAAAAACAGCAAAAACGTGAGACATTAAGAGTGAAG 4740  
 Db 996 GTCAAGTCAAGCAAAACCAAAAACAGCAAAAACGTGAGACATTAAGAGTGAAG 937  
 Qy 4741 CCGAGTGAAGATTAATTAATTAATTAATTTTAAACAGCAATTTGAAATTAATTT 4800  
 Db 936 CCGAGTGAAGATTAATTAATTAATTAATTTTAAACAGCAATTTGAAATTAATTT 877  
 Qy 4801 GTGCACTTCAGATTTCTTCAATTAATTAATTTTCAATTTTAAATTTTAAAGAAA 4860  
 Db 876 GTGCACTTCAGATTTCTTCAATTAATTAATTTTCAATTTTAAATTTTAAAGAAA 817  
 Qy 4861 TTACATATTAATTAATTAATTAATTAATTTTCAATTTTAAATTTTAACTCAATTA 4920

Db 816 TTACTATATTAATTAATTAATTAATTAATTTGAGGTAGAAATTTAATCAATTA 757  
 Qy 4921 AGGTATATTTCTTTTATTTGAGGTAGAAAGCTTTTAAAGGGATGTAAGATATCT 4980  
 Db 756 AGGTATATTTCTTTTATTTGAGGTAGAAAGCTTTTAAAGGGATGTAAGATATCT 697  
 Qy 4981 CTTCCTTTAGCTGAGAGAAAGTGAAGTTTAAATTAATTAATTAATTAATTTCCCT 5040  
 Db 696 CTTCCTTTAGCTGAGAGAAAGTGAAGTTTAAATTAATTAATTAATTAATTTCCCT 637  
 Qy 5041 GTCTTTGCTATTTGAGTTTGAACCAACAGGCGGTTGCTGAAAGGGAAACTGAAGG 5100  
 Db 636 GTCTTTGCTATTTGAGTTTGAACCAACAGGCGGTTGCTGAAAGGGAAACTGAAGG 577  
 Qy 5101 CCGGAGGAGAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5160  
 Db 576 CCGGAGGAGAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 517  
 Qy 5161 ACAAAACATTTTAGCCCAAGAAATGTCACAGAAATTTCTCAATTAACAGTATTCAGA 5220  
 Db 516 ACAAAACATTTTAGCCCAAGAAATGTCACAGAAATTTCTCAATTAACAGTATTCAGA 457  
 Qy 5221 TACAAGAAAGTTTATGATGAGTGAAGAGGAGTGAACATGATGAGTCAAGTTTCA 5280  
 Db 456 TACAAGAAAGTTTATGATGAGTGAAGAGGAGTGAACATGATGAGTCAAGTTTCA 397  
 Qy 5281 AAAGTCCAGGCTGCTGAATTTAACTTGAATCCATTCATTCAGCATTCAGACAGA 5340  
 Db 396 AAAGTCCAGGCTGCTGAATTTAACTTGAATCCATTCATTCAGCATTCAGACAGA 337  
 Qy 5341 GAGATCAGAAAGTTCAGAGATGCTCAGCTCAGCTCAAAATTTGCCAACAACAGTGTGCTACTA 5400  
 Db 336 GAGATCAGAAAGTTCAGAGATGCTCAGCTCAGCTCAAAATTTGCCAACAACAGTGTGCTACTA 277  
 Qy 5401 TACGTCAAGACCTTCAGAACCTGAGAGAGGGGAGAAACAAGTGAAGAGATGCCCA 5460  
 Db 276 TACGTCAAGACCTTCAGAACCTGAGAGAGGGGAGAAACAAGTGAAGAGATGCCCA 217  
 Qy 5461 CCTGTGAAGATGAGTGTATTAATTAATTTAGTCAATTTGATTAATTTCAATTTGCTTAA 5520  
 Db 216 CCTGTGAAGATGAGTGTATTAATTAATTTAGTCAATTTGATTAATTTCAATTTGCTTAA 157  
 Qy 5521 TCAAGAAAGCTCCGCTCTTTGCAAAATGATGAGAGAGAGTGCCTAAACTTCTA 5580  
 Db 156 TCAAGAAAGCTCCGCTCTTTGCAAAATGATGAGAGAGAGTGCCTAAACTTCTA 97  
 Qy 5581 TGTCTGATGATTTGACCTTATGCTTTAGCTCCGCTTATTAATTAATTAATTAATTAAT 5640  
 Db 96 TGTCTGATGATTTGACCTTATGCTTTAGCTCCGCTTATTAATTAATTAATTAATTAAT 37  
 Qy 5641 AGGTATTTGATTAATTTAATTAATTTATTTGTTCTCGT 5676  
 Db 36 AGGTATTTGATTAATTTAATTAATTTATTTGTTCTCGT 1  
 RESULT 5  
 AAF25906/c  
 ID AAF25906 standard; DNA; 5676 BP.  
 XX AAF25906;  
 DT 19-APR-2001 (first entry)  
 XX  
 DE Human tumor suppressor gene p51 promoter associated DNA SEQ ID 3.  
 KW Tumor suppressor; p51; cell death; cell proliferation; cancer;  
 KW cytostatic; gene therapy; screening; de.  
 XX Homo sapiens.  
 OS  
 XX  
 PN WO200100818-A1.  
 XX  
 PD 04-JAN-2001.







QY 3781 CCAAGTGTGTCAGAGACAGACGATCAAGTACCTGGGAAGCTTTAGAAATGCG 3840  
 DB 1886 CCAAGAGTGTGTCAGAGACAGACGATCAAGTACCTGGGAAGCTTTAGAAATGCG 1837  
 QY 3841 AGCTTAGGCTTCAACCCAGACCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 3900  
 DB 1836 AGCTTAGGCTTCAACCCAGACCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1777  
 QY 3901 GCTCAAGGGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 3960  
 DB 1776 GCTCAAGGGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1717  
 QY 3961 ATATTAATGGAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 4020  
 DB 1716 ATATTAATGGAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1657  
 QY 4021 CCTTCTATTTAGTCAATCAATTTCTAATGAGGGAAGAAAGCCAACTCTTAACT 4080  
 DB 1656 CCTTCTATTTAGTCAATCAATTTCTAATGAGGGAAGAAAGCCAACTCTTAACT 1597  
 QY 4081 GCAATTAACAAATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4140  
 DB 1596 GCAATTAACAAATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1537  
 QY 4141 TGAAGCAAGCTTTTCAATTTGCTGAGGCTCAAGCTGAGGCTTTTGAATTCACAA 4200  
 DB 1536 TGAAGCAAGCTTTTCAATTTGCTGAGGCTCAAGCTGAGGCTTTTGAATTCACAA 1477  
 QY 4201 CACAGATGTTAGCGGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 4260  
 DB 1476 CACAGATGTTAGCGGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1417  
 QY 4261 TCTCCCTGATGCTGAGGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 4320  
 DB 1416 TCTCCCTGATGCTGAGGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1357  
 QY 4321 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4380  
 DB 1356 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1297  
 QY 4381 TTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4440  
 DB 1296 TTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1237  
 QY 4441 AGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4500  
 DB 1236 AGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1177  
 QY 4501 TTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4560  
 DB 1176 TTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1117  
 QY 4561 ACCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4620  
 DB 1116 ACCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1057  
 QY 4621 TTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4680  
 DB 1056 TTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 997  
 QY 4681 GTCAAGTCAAGGGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 4740  
 DB 996 GTCAAGTCAAGGGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 937  
 QY 4741 CCGAGTCAAGGGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 4800  
 DB 936 CCGAGTCAAGGGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 877  
 QY 4801 GTGAGTCAAGGGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 4860  
 DB 876 GTGAGTCAAGGGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 817  
 QY 4861 TTACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4920

DB 816 TTACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 757  
 QY 4921 AGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4980  
 DB 756 AGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 697  
 QY 4981 CTTTCTATTTAGTCAATCAATTTCTAATGAGGGAAGAAAGCCAACTCTTAACT 5040  
 DB 696 CTTTCTATTTAGTCAATCAATTTCTAATGAGGGAAGAAAGCCAACTCTTAACT 637  
 QY 5041 GTCTTGTATTTAGTCAATCAATTTCTAATGAGGGAAGAAAGCCAACTCTTAACT 5100  
 DB 636 GTCTTGTATTTAGTCAATCAATTTCTAATGAGGGAAGAAAGCCAACTCTTAACT 577  
 QY 5101 CCGGGAAGGGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5160  
 DB 576 CCGGGAAGGGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 517  
 QY 5161 ACATAACATTTTACCCGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5220  
 DB 516 ACATAACATTTTACCCGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 457  
 QY 5221 TACAAGAGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5280  
 DB 456 TACAAGAGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 397  
 QY 5281 AAAGTCAAGGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5340  
 DB 396 AAAGTCAAGGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 337  
 QY 5341 GAGATCAAGGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5400  
 DB 336 GAGATCAAGGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 277  
 QY 5401 TACGTCAAGGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5460  
 DB 276 TACGTCAAGGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 217  
 QY 5461 GCTGTCAAGGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5520  
 DB 216 GCTGTCAAGGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 157  
 QY 5521 TCAAGAAAGCTGCTGCTTCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5580  
 DB 156 TCAAGAAAGCTGCTGCTTCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 97  
 QY 5581 TGTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5640  
 DB 96 TGTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 37  
 QY 5641 AGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5676  
 DB 36 AGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1  
 RESULT 6  
 ABX5655/C  
 ID ABX5655 standard; DNA; 118384 BP.  
 XX ABX5655:  
 XX  
 DT 20-FEB-2003 (first entry)  
 XX  
 XX Human autoimmune disease related protein PRT1 gene region #1.  
 XX  
 XX PRT1; human; autoimmune disease; psoriasis; type I diabetes;  
 KM Rheumatoid arthritis; cation-chloride transport; gene therapy; gene; db.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO20028042-A2.





```

FT      /*tag= j
FT      /number= 5
FT      intron 28396..28902
FT      /*tag= k
FT      exon 28903..28975
FT      /*tag= l
FT      /number= 6
FT      intron 28976..35372
FT      /*tag= m
FT      exon 35373..35442
FT      /*tag= n
FT      /number= 7
FT      intron 35443..35705
FT      /*tag= o
FT      exon 35706..35761
FT      /*tag= p
FT      /number= 8
FT      intron 35762..36266
FT      /*tag= q
FT      exon 36267..36359
FT      /*tag= r
FT      /number= 9
FT      intron 36360..36591
FT      /*tag= s
FT      exon 36592..36684
FT      /*tag= t
FT      /number= 10
FT      intron 36685..38529
FT      /*tag= u
FT      exon 38530..38672
FT      /*tag= v
FT      /number= 11
FT      intron 38673..39376
FT      /*tag= w
FT      exon 39377..39562
FT      /*tag= x
FT      /number= 12
FT      intron 39563..40050
FT      /*tag= y
FT      exon 40051..40129
FT      /*tag= z
FT      intron 40130..45672
FT      /*tag= aa
FT      exon 45673..47158
FT      /*tag= ab
FT      /number= 14
FT
FT
XX      WO200292629-A1.
XX
XX      21-NOV-2002.
XX
XX      14-MAY-2002; 2002WO-AU000591.
XX
XX      14-MAY-2001; 2001AU-00004971.
XX
XX      (UTOU-) UNIV QUEENSLAND TECHNOLOGY.
XX
XX      Clements JA;
XX
XX      WPI; 2003-129264/12.
XX      P-PSDB; AAE34613.
XX
XX      New human tweety homolog 2 polypeptides and polynucleotides, useful for
XX      producing an antigen-binding molecule that is immuno-interactive with the
XX      polypeptide or as diagnostic markers for cancers.
XX
XX      Claim 10; Page 128-156; 176pp; English.
XX
XX      The invention relates to human tweety homologue 2 (TTYH2) polypeptide and
XX      polynucleotide sequence. TTYH2 is useful for producing an antigen-binding
XX      molecule that is immuno-interactive with the polypeptide. The agent is
XX      useful for manufacturing a medicament for restoring a normal level and/or

```

```

CC      functional activity of TTYH2 expression in a patient, and for treating or
CC      preventing cancer or tumour. TTYH2 sequences may also be used to provide
CC      both drug targets and regulators to promote or inhibit one or more
CC      activities, and to provide diagnostic markers for cancers. The present
CC      sequence is human TTYH2 gene
CC
XX      SQ      Sequence 47999 BP; 11178 A; 12727 C; 12874 G; 11209 T; 0 U; 11 Other;
XX
XX      Query Match      3.1%; Score 185.4; DB 7; Length 47999;
XX      Best Local Similarity 74.9%; Pred. No. 3.6e-25;
XX      Matches 245; Conservative 0; Mismatches 81; Indels 1; Gaps 1;
XX
QY      3284 AGATGAGTGGTGGCTGATGCTGATATCCCAACACTTGGAGAGCCGGCTGGAGAGA 3343
DB      18395 AGTGGAGTGGTGGCTGACGCTTAATCCAGCACTTGGAGAGCCCAAGTGGAGAGA 18454
QY      3344 TAGCTGAGTCCAGAGTTTGAGACCACTGAGGCAACAGCAAGACCCATATCTTAA 3403
DB      18455 TTGCTGAGGCGCAGGAGTTCAAGACCACTGGGCAACATAGCAAGACCTATCTTACA 18514
QY      3404 AAACAAACAAACAAATTAATCTGGGTATGGTGTGCTACCTGATGCCAAGCTACA 3463
DB      18515 AAAATTAAAAAATAAATAAAGTACGTAGTGTGAGTGTGACCTGTGTCCCAACACT 18574
QY      3464 CAGAAAGCTGAGCGCAAGATCACTTGAAGCCAGAGAGTTGAGGCTGCATGATCAATG 3523
DB      18575 CAGAGAGCTGAGGTGGAGAGATTGCTTGAAGCCAGAGATTGAGGCTGCATGATG 18634
QY      3524 AACGCGTCTCACT-CAGTCTGGGTGACAGTGCAGAGAAAGCTGTCTCAAAAAATAATAA 3582
DB      18635 ATCAGACACTGCGCTCCAGCTTGGGTGACAGAGTGAAGACTTGTCTAAAAAAG 18694
QY      3583 TAAATTAAATTAATCTTTAAAAAACAA 3609
DB      18695 AAAAGAAAAAAGAAAAAGAAATGATCA 18721
DB
DB      RESULT 10
DB      ADCC4703/C
DB      ID      ADCC4703 standard; DNA; 335199 BP.
XX
XX      AC      ADCC4703;
XX
XX      DT      18-DEC-2003 (first entry)
XX
XX      DE      Human wild type HML4Y genomic sequence.
XX
XX      KW      de; nootropic; neuroleptic; tranquilizer; gene therapy; synaptogenesis;
XX      mutation; neurological disease; mental disorder; psychiatric illness;
XX      autism; Asperger syndrome; schizophrenia;
XX      attention deficit hyperactivity disorder.
XX
XX      OS      Homo sapiens.
XX
XX      PH      Key
XX      FT      exon 10156..10298
XX      /tag= a
XX      /number= 1b
XX      FT      intron 10299..108499
XX      /tag= b
XX      /number= 1b
XX      FT      exon 108500..109001
XX      /tag= c
XX      /number= 2
XX      FT      intron 109002..205868
XX      /tag= d
XX      /number= 2
XX      FT      exon 205869..205928
XX      /tag= e
XX      /number= 2b
XX      FT      intron 205929..209525
XX      /tag= f
XX      /number= 2b

```

FT exon 209526..209679  
 FT /\*cag= 9  
 FT /number= 3  
 FT intron 209680..235027  
 FT /\*cag= h  
 FT /number= 3  
 FT exon 235028..235139  
 FT /\*cag= 1  
 FT /number= 3b  
 FT intron 235140..238089  
 FT /\*cag= j  
 FT /number= 3b  
 FT exon 238090..238212  
 FT /\*cag= k  
 FT /number= 3c  
 FT intron 238213..310596  
 FT /\*cag= 1  
 FT /number= 3c  
 FT exon 310597..310783  
 FT /\*cag= m  
 FT /number= 4  
 FT intron 310784..316138  
 FT /\*cag= n  
 FT /number= 4  
 FT exon 316139..316929  
 FT /\*cag= o  
 FT /number= 5  
 FT intron 316930..326821  
 FT /\*cag= p  
 FT /number= 5  
 FT exon 326822..330136  
 FT /\*cag= q  
 FT /number= 6  
 PN MO2003045998-A2.  
 XX 05-JUN-2003:  
 PD 02-DEC-2002; 2002MO-FR004134.  
 PF 30-NOV-2001; 2001CA-02364106.  
 PR (INRM) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (INSP) INST PASTEUR  
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITALUX PARIS.  
 XX Bourgeron T, Jamain S, Quach H, Betancur C, Leboyer M,  
 PI Gillberg C;  
 XX WPI; 2003-493399/46.  
 DR New nucleic acid encoding mutant protein involved in synaptogenesis,  
 XX useful for treatment and diagnosis of e.g. autism, Asperger syndrome, and  
 PT schizophrenia.  
 PT schizophrénia.  
 XX  
 PS Claim 10; SEQ ID NO 4; 416pp: French.  
 XX  
 CC The invention relates to an isolated or purified polynucleotide encoding  
 CC a polypeptide (the wild-type form of which is involved in synaptogenesis)  
 CC that includes at least one mutation associated with development of  
 CC neurological disease and/or a predisposition to development of mental  
 CC disorders or psychiatric illness. The polypeptide are used to screen for  
 CC agents that modulate their activity. Also nucleic acid, polypeptide,  
 CC polypeptide-specific antibodies, vectors containing he nucleic acid and  
 CC host cells containing the vector, are useful as pharmaceuticals for  
 CC treating mental and neurological disorders, specifically autism, Asperger  
 CC syndrome, schizophrenia and attention deficit hyperactivity disorder. The  
 CC wild-type forms of the nucleic acid and polypeptide can be used  
 CC similarly. Also detecting mutations in the nucleic acid and polypeptide,  
 CC or measuring activity of the polypeptide, can be used to detect  
 CC biochemical disorders that affect formation of synapses and to diagnose  
 CC mental disease. This sequence corresponds to the genomic sequence of the  
 CC human wild type HNLAY gene.

XX SQ Sequence 335199 BP; 98958 A; 63589 C; 65677 G; 106975 T; 0 U; 0 Other;  
 Query Match 3.1%; Score 183; DB 9; Length 335199;  
 Best Local Similarity 65.2%; Pred. No. 1.3e-24;  
 Matches 303; Conservative 0; Mismatches 155; Indels 7; Gaps 2;  
 QY 3158 TTTCGTCTCTCTTACTAGCTTATGATTTTAAATPACATATGAAATCATG 3217  
 Db 239974 TTTATATATCTTATTAAGATCTTTATATATATATATATATATATATAT 239915  
 QY 3218 ACATACGTTTAAATGATATATATATATATATATATATATATATATATAT 3277  
 Db 239914 ATGTTACAAAATGTATATATATATATATATATATATATATATATATAT 239855  
 QY 3278 AAAATAGATGATGCTGTGCTGCTCATGCTGTATATCCCACTTTGGAAAGCCGAGTGG 3337  
 Db 239854 AAAATGCGCCAGGTCCGGATGCTCATGCTGTATATCCAGCATTTGGAGGCCAGGTG 239795  
 QY 3338 GAGGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 3397  
 Db 239794 GGTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 239735  
 QY 3398 TCTAAAAAACAACAAAACAAAATTAACCTGATGCTGATGCTGATGCTGATGCTGAT 3457  
 Db 239734 TCTATGGAATACAA-----AAATTAAGTGAAGCTGTGTGCTGATGCTGATGCTGAT 239681  
 QY 3458 GCTACACAGAGAGCTGAGCAGAGATCACTTGAAGCCAGAGAGCTGAGCTGAGCTGAG 3517  
 Db 239680 GCTACTGGGAGAGCTGAGCAGAGATCACTTGAAGCCAGAGAGCTGAGCTGAGCTGAG 239621  
 QY 3518 TCCATGAAACCGCTGCTAAGCT-CAGTCTGGGTGACAGTGCAGAGAGCTGCTCAAAAT 3576  
 Db 239620 CCTAGATCAACGCCATTCAGCTCCAGCTGGGTGACAGAGAGCTGCTGCTCAAAAAA 239561  
 QY 3577 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3621  
 Db 239560 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 239516  
 RESULT 11  
 AAD28072  
 ID AAD28072 standard; DNA; 36651 BP.  
 XX  
 AC AAD28072;  
 XX  
 DT 22-APR-2002 (first entry)  
 XX  
 DE Human kinase genomic DNA.  
 XX  
 KM Human; kinase; G-protein coupled receptor kinase; tissue differentiation;  
 KM chromosome 3; therapeutic; immune response; drug screening; enzyme; SNP;  
 KM single nucleotide polymorphism; de.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT replacement(2022,T)  
 FT /\*cag= a  
 FT /standard name= "Single nucleotide polymorphism"  
 FT CDS 2076..33528  
 FT /\*cag= b  
 FT /product= "Human kinase protein"  
 FT exon 2076..2687  
 FT /\*cag= c  
 FT intron 2688..4164  
 FT /\*cag= d  
 FT variation replacement(3672,T)  
 FT /\*cag= e  
 FT /standard name= "Single nucleotide polymorphism"  
 FT FT variation replacement(3884,C)  
 FT /\*cag= f  
 FT /standard\_name= "Single nucleotide polymorphism"

```

FT      exon
FT      4165..4602
FT      /tag= g
FT      intron
FT      4603..31231
FT      /tag= h
FT      variation
FT      replace(4986,T)
FT      /tag= i
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(5268,C)
FT      /tag= j
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(7344,C)
FT      /tag= k
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(7414,C)
FT      /tag= l
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(8113,A)
FT      /tag= m
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(8394,T)
FT      /tag= n
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(9233,A)
FT      /tag= o
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(9255,T)
FT      /tag= p
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(9746..9747,A)
FT      /tag= q
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(9747,M)
FT      /tag= x
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(11634..11635,C)
FT      /tag= s
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(11635,A)
FT      /tag= t
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(13721,T)
FT      /tag= u
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(13884,C)
FT      /tag= v
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(13987,A)
FT      /tag= w
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(14376..14377,C)
FT      /tag= x
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(15045,T)
FT      /tag= y
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(15742,G)
FT      /tag= z
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(17482,G)
FT      /tag= aa
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(17688,G)
FT      /tag= ab
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(17992,C)
FT      /tag= ac
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(19132..19133,G)
FT      /tag= ad
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(20105,G)
FT      /tag= ae
FT      /standard_name= "Single nucleotide polymorphism"

```

```

FT      variation
FT      replace(24194..24195,T)
FT      /tag= af
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(24219..24220,C)
FT      /tag= ag
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(24220,A)
FT      /tag= ah
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(24278,A)
FT      /tag= ai
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(30132,C)
FT      /tag= aj
FT      /standard_name= "Single nucleotide polymorphism"
FT      exon
FT      31232..31506
FT      /tag= ak
FT      intron
FT      31507..33591
FT      /tag= al
FT      exon
FT      33592..33928
FT      /tag= am
FT      variation
FT      replace(33800,M)
FT      /tag= an
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(34513,A)
FT      /tag= ao
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(34821,G)
FT      /tag= ap
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(34983,A)
FT      /tag= aq
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(35153,T)
FT      /tag= ar
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(36368,T)
FT      /tag= as
FT      /standard_name= "Single nucleotide polymorphism"
FT      PN
FT      MO200192496-A2.
FT      XX
FT      PD
FT      06-DEC-2001.
FT      XX
FT      31-MAR-2001; 2001WO-US017510.
FT      PF
FT      01-JUN-2000; 2000US-0206331P.
FT      PR
FT      18-DEC-2000; 2000US-00738894.
FT      XX
FT      (APPL-) APPLERA CORP.
FT      PA
FT      Guegler K, DI Francesco V, Beasley EM;
FT      PI
FT      WPI: 2002-130533/17.
FT      DR
FT      P-PSDS; AAB17136.
FT      DR
FT      XX
FT      New isolated human kinase proteins and nucleic acids, useful as a major
FT      PT target for drug action and development, particularly for screening
FT      modulators of the kinase peptides.
FT      XX
FT      Claim 4; Fig 3; 87pp; English.
FT      XX
FT      The invention relates to human kinase proteins that are related to G-
FT      CC protein coupled receptor kinase subfamily. Human kinase gene is located
FT      CC on chromosome 3. The kinase peptide and nucleic acid are useful in the
FT      CC development of human therapeutic and diagnostic compositions. The peptide
FT      CC is useful as a target for drug action and development. The peptide
FT      CC is valuable to the field of pharmaceutical development to identify and
FT      CC characterize modulators of the kinase. The proteins may also be used to
FT      CC raise antibodies or to elicit an immune response, as a reagent in assays
FT      CC designed to quantitatively determine levels of the protein in biological
FT      CC fluids and as markers for tissues in which the corresponding protein is
FT      CC preferentially expressed (either constitutively or at a particular stage

```





FT	/tag= b
FT	/strand_name= "Single nucleotide polymorphism"
FT	CDS
FT	3000..21730
FT	/tag= c
FT	/product= "human secreted protein"
FT	/note= "contains 2 introns"
FT	3000..3087
FT	exon
FT	/tag= d
FT	/number= 1
FT	3088..7777
FT	/tag= e
FT	/number= 1
FT	replace(5184..5185,C)
FT	/tag= f
FT	/strand_name= "Single nucleotide polymorphism"
FT	7778..7909
FT	/tag= g
FT	/number= 2
FT	7910..21575
FT	/tag= h
FT	/number= 2
FT	replace(8753,G)
FT	/tag= i
FT	/strand_name= "Single nucleotide polymorphism"
FT	variation
FT	replace(14860,C)
FT	/tag= j
FT	/strand_name= "Single nucleotide polymorphism"
FT	replace(15865,A)
FT	/tag= k
FT	/strand_name= "Single nucleotide polymorphism"
FT	variation
FT	replace(16060,G)
FT	/tag= l
FT	/strand_name= "Single nucleotide polymorphism"
FT	variation
FT	replace(16510,C)
FT	/tag= m
FT	/strand_name= "Single nucleotide polymorphism"
FT	variation
FT	replace(19213,T)
FT	/tag= n
FT	/strand_name= "Single nucleotide polymorphism"
FT	variation
FT	replace(20396,C)
FT	/tag= o
FT	/strand_name= "Single nucleotide polymorphism"
FT	exon
FT	21576..21727
FT	/tag= p
FT	/number= 3
FT	variation
FT	replace(21904,G)
FT	/tag= q
FT	/strand_name= "Single nucleotide polymorphism"
FT	variation
FT	replace(22425,G)
FT	/tag= r
FT	/strand_name= "Single nucleotide polymorphism"
FT	variation
FT	replace(22575,G)
FT	/tag= s
FT	/strand_name= "Single nucleotide polymorphism"
PX	MO2003006481-A2.
PX	PN
PX	PD
PX	23-JAN-2003.
PX	PF
PX	10-JUL-2002; 2002MO-USO21574.
PX	PR
PX	10-JUL-2003; 2003US-00901152.
PA	(APPL-) APPLERA CORP.
PX	PI
PX	Hu S, Ketchum K, Ladunga I;
XX	WPI: 2003-210419/20.
DR	P-PDB: ABBY1736.
PT	New isolated peptide, useful for treating a disease or condition mediated by a human secreted protein e.g. cancer.
XX	

PS Claim 1 b; Fig 3; 93pp; English.

CC The invention relates to a novel human secreted protein that is related  
CC to the epidermal growth factor (EGF) superfamily. The activity of the  
CC protein of the invention may be described as cytoskeletal. The protein of  
CC the invention is useful for preparing a composition for diagnosing or  
CC treating a disease or condition mediated by human secreted proteins,  
CC including cancer. The gene encoding the protein of the invention may be  
CC found on human chromosome 22. The current sequence represents the human  
CC secreted protein genomic DNA sequence

SQ Sequence 58985 BP; 13362 A; 14853 C; 15805 G; 13968 T; 0 U; 997 Other;

Query Match	3.04	Score 181;	DB 8;	Length 58985;
Best Local Similarity	70.84;	Pred. No. 2.5e-24;		
Matches 271; Conservative	0;	Mismatches 105;	Indels 7;	Gaps 22;

Oy 3250 TACGCTATTAACCTTTTAAATTTTAAAAAAATAGATAGTGTGCTGCTACGCTGT 3309  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 8142 TGCGCTTGAACCTTTTGTCTTTTAAAGAAATGCGTGGTGGTGCCTACAGTCTGT 8201

```

Oy      3310  AATCCCAACACTTTGGGAAGCCGGGTCGGGAGCATAGCTTGATGTCACAGCATGTTGAGACC 3366
      |||||
Db      8202  AATCCCAAGCACTTGGGAGACCGAGGCAAGCAAGTACCTGAGGCTCAAGCATGTTGAGACC 8266

```

```
QY      3370 AGTACGGGCAACAGCAGACCCTATCT-----AAATAAACAAAACAAAAT 3422
          ||| | | | | | | | | | | | | | | | | | | | | | | | | |
DB      8262 AGCTGGCCACATCGCAAACCCCATCTACTGAATAACACAAATACAAAAAAGT 8323
```

Oy 3424 TACCTGGATAGTGTGTCACCTGTACTGCCAAGCTACACAGAAGCTGAGCGAGAGC 3488  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 8322 TAGCCGAGCGTGCTGCATGCTATAATTCACTACTCGTGAAGCTGAGGCGAGAGA 8386

Oy 3484 ATCACTTGAGCCCGACGAGGTGGAGCTGCAGTATCATTAACGCCGCTGTACAAT - CAG 3544  
Db 8382 ATCACTTGAACCCGACGAGCAGAGCTTGCACTGAGCCGAGATGGCGCACACTGCACATCCAG 8444

Oy 3543 TCTGGGTGACAGTCAAGACGCTGCTCAAAAATATATAATATAAACTTTAA 360  
 |||||  
 Db 8442 CTTGGGCGACAGACAGACTCTGCTCAAAAAAAAAAGAAAGAAAGAAAGAAATTAA 850

Oy	3603	AAACAAATTAATTAATTTT	3625
Db	8502	GATGAGCATTGATGAGTATT	8524

RESULT 14  
ABZ22654

XX  
AC  
XX  
ABZ22654;  
XX

XX Human epidermal growth factor related secreted protein DNA SEQ ID NO:3  
DB  
XX

KM epidermal growth factor; EGF; gene therapy; therapeutic; drug screening;  
KM biological activity; immune response; chromosome 22; SNP; gene;  
KM single nucleotide polymorphism; ds.

OS	Homo sapiens.	Location/Qualifiers
XX		
EH	Key	
EM	variations	location(1277-8)

```

FT      /*tag= a
FT      /standard name= "single nucleotide polymorphism (SNP) "
FT      2076. .140828
CDS
FT      /*tag= b

```

FT	/product= "EGF protein subfamily related secreted protein"
FT	/note= "contains introns"
FT	2076..2163
exon	

```
FT FT /*tag= c  
FT /number= 1  
FT intron 2164. .6058  
FT /*tag= d  
FT /number= 1  
FT variation replace(4260. .4261,C)  
FT /*tag= e  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(4814,C)  
FT /*tag= f  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT 6059. .6190  
FT /*tag= g  
FT /number= 2  
FT intron 6191. .25226  
FT /*tag= h  
FT /number= 2  
FT variation replace(7034,A)  
FT /*tag= i  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(8334. .8335,G)  
FT /*tag= j  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(13141,G)  
FT /*tag= k  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(14146,G)  
FT /*tag= l  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(14341,G)  
FT /*tag= m  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(14791,C)  
FT /*tag= n  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT 15626. .115742  
FT /*tag= et  
FT /number= 9  
FT variation replace(17494,T)  
FT /*tag= o  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(18678,C)  
FT /*tag= p  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(20187,A)  
FT /*tag= q  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(20708,G)  
FT /*tag= r  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(20858,G)  
FT /*tag= s  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(21858,A)  
FT /*tag= t  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(24101,A)  
FT /*tag= u  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(24234,T)  
FT /*tag= v  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(24974,A)  
FT /*tag= w  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT 25227. .25355  
FT /*tag= x  
FT /number= 3  
FT intron 25356. .54132  
FT /*tag= y  
FT /number= 3  
FT variation replace(25436,A)  
FT /*tag= z  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(25450,C)  
FT /*tag= aa  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(25733. .24734,T)  
FT /*tag= ab  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(26583,C)  
FT /*tag= ac  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(26966,A)  
FT /*tag= ad  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(26980,G)  
FT /*tag= ae  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(26984,C)  
FT /*tag= af  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(27044,A)  
FT /*tag= ag  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(28933,G)  
FT /*tag= ah  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(30219,C)  
FT /*tag= ai  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(30949,G)  
FT /*tag= aj  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(32377,C)  
FT /*tag= ak  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(32395,G)  
FT /*tag= al  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(34305,A)  
FT /*tag= am  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(34538,R)  
FT /*tag= an  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(35369,G)  
FT /*tag= ao  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(35620,T)  
FT /*tag= ap  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(40078,C)  
FT /*tag= aq  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(42292,C)  
FT /*tag= ar  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(43765,A)  
FT /*tag= as  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(44837,A)  
FT /*tag= at  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(44995. .44996,A)  
FT /*tag= au  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(45363,T)  
FT /*tag= av  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(46514,Y)  
FT /*tag= ax  
FT /standard_name= "single nucleotide polymorphism (SNP)"  
FT replace(52639,C)  
FT /*tag= ay  
FT /standard_name= "single nucleotide polymorphism (SNP)"
```



PR 14-SEP-2000; 2000US-0231064P.  
 PR 14-SEP-2000; 2000US-0231065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235835P.  
 PR 29-SEP-2000; 2000US-0236127P.  
 PR 29-SEP-2000; 2000US-0236167P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 02-OCT-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241222P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249246P.  
 PR 17-NOV-2000; 2000US-0249255P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 01-DEC-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI, 2001-465566/50.  
 XX  
 PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,  
 PT treating neural, immune system, muscular, reproductive, pulmonary,  
 PT cardiovascular, renal, proliferative disorders and cancerous diseases.  
 XX  
 PS Disclosure; SEQ ID NO 2167; 1180bp; English.  
 XX  
 CC The present invention relates to the isolation of novel human enzyme  
 CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences  
 CC encoding them. The enzyme polypeptides of the invention may comprise the  
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,  
 CC isomerases or ligases. The sequences of the invention are useful in the  
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of  
 CC disorders including hyperproliferative disorders (e.g. cancer),  
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.  
 CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic  
 CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),  
 CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders  
 CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and  
 CC infectious disorders (e.g. influenza). The polynucleotides of the  
 CC invention can also be used in gene therapy. MS4185-MS42192 represent  
 CC DNA sequences encoding for the novel human enzyme polypeptides of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WFO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 26329 BP; 6954 A; 5487 C; 5702 G; 8186 T; 0 U; 0 Other;

Query Match 3.0%; Score 180; DB 4; Length 26329;  
 Best Local Similarity 71.6%; Pred. No. 3.5e-24;  
 Matches 252; Conservative 0; Mismatches 95; Indels 5; Gaps 1;

QY 3263 TTTTAATTTTAAAAAATATAGATGATGTCGTCGTCATGCTGTAATCCCAACTT 3322  
 DB 5202 TTATATAGAGCTACAGTAATTCAGAGCTGTGTGATTCGCTCAGGCTGTATCCCAACTT 5143  
 QY 3323 TGGAGAGCCGAGCTCGAGAGATAGCTTGAAGTCAGCACTTGTAGACAGAGCA 3382  
 DB 5142 TGGAGAGCCGAGAGATATTAATCTTGAAGTCAGCACTTGTAGACAGAGCA 5083  
 QY 3383 CAGCAAGCCGATCTTAAAAAACAACAAATTAACCTGGTATGCTTGC 3442  
 DB 5082 TAGTGAGACCTATCTTCA-----AAAAGATAAAAATTAACCTGAAGCAGTGTGC 5028  
 QY 3443 TCACCTGATGTCAGATACAGAGCTGAGAGAGATCACTTGAAGCCAGAGAG 3502  
 DB 5027 ACGCTGATGTCAGATACAGAGCTGAGAGAGATCACTTGAAGCCAGAGAG 4968  
 QY 3503 TTGAGAGCTGAGATATCAAGAGAGAGCTGATCACTGATGCTGAGAGAGAGAG 3562  
 DB 4967 CCAAGCTGAGATATCAATGATGATGATGATGATGATGATGATGATGATGATG 4908  
 QY 3563 GCTGCTCAAAAT 3614  
 DB 4907 CTTGCTCAAAAT 4856

Search completed: September 27, 2004, 08:20:51  
 Job time : 2046 secs



Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Sequence Tagged Connector  
 Plate: 3165 row: L column: 16  
 Class: BAC ends  
 High quality sequence stop: 523.  
 Location/Qualifiers

FEATURES  
 SOURCE

1. 523  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=3165 Col=16 Row=L"  
 /sex="male"  
 /clone\_lib="CIR Approved Human Genomic Sperm Library D"  
 /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in  
 E-Coli DH10B"

## ORIGIN

Query Match 7.2% Score 427.8 DB 28 Length 523:  
 Best Local Similarity 90.6% Pred No. 6.9e-52; Indels 0; Gaps 0;

Matches 453; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 672 GATCTCACTGACAGGCTGCGAGCTCTACGGAACATCAATCTCTTAAAC 731  
 DB 12 GATCTCACTGACAGGCTGCGAGCTCTACGGAACATCAATCTCTTAAAC 71  
 QY 732 CTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 791  
 DB 72 CTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 131  
 QY 792 ACGAAGCTTATTTCTCTGAAAGCTCTGATGATGATGATGATGATGATGATGAT 851  
 DB 132 ACGAAGCTTATTTCTCTGAAAGCTCTGATGATGATGATGATGATGATGATGAT 191  
 QY 852 CATGATGAGGCTCTCTGCGAACAATGCAATCTTTCTTTGGAATTTGAATCTTAAG 911  
 DB 192 CATGATGAGGCTCTCTGCGAACAATGCAATCTTTCTTTGGAATTTGAATCTTAAG 251  
 QY 912 CTGAATGCTGAAGTTCAAAAAAGCTGTTGATGATGATGATGATGATGATGATGAT 971  
 DB 252 CTGAATGCTGAAGTTCAAAAAAGCTGTTGATGATGATGATGATGATGATGATGAT 311  
 QY 972 AAAGTGAAGCTGATCTCTTCTTAAGTCCCTGAATTTAATTTATCTTCTGTTACG 1031  
 DB 312 AAAGTGAAGCTGATCTCTTCTTAAGTCCCTGAATTTAATTTATCTTCTGTTACG 371  
 QY 1032 CCTTTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1091  
 DB 372 CCTTTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 431  
 QY 1092 TTTCTTTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 1151  
 DB 432 TTTCTTTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 491  
 QY 1152 GCGAGATCAAGTTCTGTTG 1171  
 DB 492 GAAAGAAATGATCTATTCTG 511

RESULT 2  
 BY723946 638 bp mRNA linear EST 17-DEC-2002  
 LOCUS BY723946 RIKEN full-length enriched adult male hypothalamus Mus  
 DEFINITION musculus cDNA clone A230061p11 5', mRNA sequence.  
 ACCESSION BY723946  
 VERSION BY723946.1 GI:27137063  
 KEYWORDS EST,  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

REFERENCE  
 1 (bases 1 to 638)  
 AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT  
 12466851  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 Tel.: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gs.c.riken.go.jp,  
 url: http://genome.gsc.riken.go.jp/  
 Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P.,  
 Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,  
 Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,  
 Kondo, S., Komuro, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, K.,  
 Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,  
 Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tegan, M.,  
 Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M., and Hayashizaki, Y.  
 Direct Submission  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system—384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.

## FEATURES

## SOURCE

1. 638  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="A230061p11"  
 /sex="male"







REFERENCE 1 (bases 1 to 567)  
 AUTHORS Robertson,N.G., Kleierpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R. and Morton,C.C.  
 TITLE Isolation of novel and known genes from a human fetal cochlear cDNA library using subtractive hybridization and differential screening  
 JOURNAL Genome 23, 42-50 (1994)  
 MEDLINE 9513011  
 PUBMED 7829101  
 COMMENT Contact: Morton, C. C.  
 Department of Pathology and Obstetrics, Gynecology and Reproductive Biology  
 Brigham and Women's Hospital  
 75 Francis Street, Harvard Medical School, Boston, MA 02115, USA  
 Tel: 617 732 7880  
 Fax: 617 736 6396  
 Email: comorton@rics.bwh.harvard.edu  
 DNA sequencing and analyses were performed by National Institutes of Health Intramural Sequencing Center (NISC; see <http://www.nisc.nih.gov>).  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
 Plate: LLM6320 row: F column: 10  
 Seq primer: M13RP1 reverse primer (ABI).  
 Location/Qualifiers  
 1..567  
 /organism="Homo sapiens"  
 /mol\_type="RNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2537481"  
 /tissue\_type="cochlea"  
 /dev\_stage="16-22 week fetus"  
 /lab\_host="SOLR cells (kanamycin resistant)"  
 /clone\_lib="Morton Fetal Cochlea"  
 /note="Organ: ear; Vector: pBluescript SK-; Site: 1: EcoRI; Site: 2: XhoI; Reference: Genom 23, 42-50 (1994) Cloned unidirectionally. Primer: Oligo dt. Fetal cochlea, normal. 37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-2AP XR Vector. Library constructed by N. Robertson, C. Morton. -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3' "

ORIGIN  
 Query Match 3.1%; Score 184.6; DB 12; Length 567;  
 Best Local Similarity 68.6%; Pred. No. 4.8e-17;  
 Matches 300; Conservative 0; Mismatches 129; Indels 8; Gaps 3;

QY 3170 TTAGCTAGTACTTAGGATTTTAAATACAACTATGAAATCATGACATACGTTAA 3229  
 DB 59 TTGCTTAAAGTTTTCATATTTGAAAAAACTCCACATCAGTACATGATCCTC 118  
 QY 3230 ATGATATTATTAAATAGGCTATTAACCTTTAAATTTTAAAAAAATGATGA 3289  
 DB 119 ATATTTTCTTTTGTATCTTTGATTTTA---TTAAATTTTATTAATATAGCTCG 175  
 QY 3290 GTGTGTGTGCTCATGCTGTATTCCTCAACACTTTGGGAGCGCGGTGGAGATAGCTT 3349  
 DB 176 GTGAGGTGCTTACTCTGTATTCCTCAACACTTTGAGAGGCTGAGGTGAGATATTCCT 235  
 QY 3350 GAGTCAGCAGTTTGAACAGTCAAGGCAACAGAGCAAGCCCATATCTAAAAAACA 3409  
 DB 216 GAGCTCAGAGTTCAGACACCCCTGGGCAACAGAGCAAGTACTTCTCAAAAAA 295  
 QY 3410 AAACAAACAAATTAATCTGGGATATGATGTGCTCACTGTAGTCAAGTACACAGAA 3469  
 DB 296 ATAGA-----AAATTAAGTGGGCAATGAGGCTGTGCTGTAGTCCAGTACTCGGAG 351  
 QY 3470 GCTAAGCAGAGATCATTTAGCTCCAGAGGTTTGAAGCTGCAATGATCAAGCGG 3529  
 DB 352 GCTAAGGCGGGGATCATTTAGCTCCAGAGGTTTGAAGCTGCAATGATCAAGCGG 411  
 QY 3530 CTGCTACACT-CAGTCTGGGTGACAGTCAAGAGCTGTCTCAAAAATATATATATATA 3588  
 DB 412 CTGCTGCACTCCAGCGTGGGTGACAGAGCAAGACCTGTCTCAAAAACCCCAAAAAACA 471

QY 3589 AAATTAACCTTTAAAAA 3605  
 DB 472 AACGAAATATCAAAATA 468

RESULT 7  
 A0275866/c 664 bp DNA linear GSS 22-NOV-1998  
 LOCUS CITBI-El-2517A7.TF CITBI-El Homo sapiens genomic clone 2517A7,  
 DEFINITION genomic survey sequence.  
 ACCESSION A0275866  
 VERSION A0275866.1 GI:3901682  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Buxarjota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 664)  
 Adams,M.D., Rounsley,S.D., Zhao,S., Baas,S., Linber,K., Golden,K.,  
 Berry,K., Granger,D., Sub,E., Wible,C., Shizuya,H., Simon,M. and  
 Venter,J.C.  
 Use of a random human BAC End Sequence Database for Sequence-ready  
 Map Building  
 Unpublished (1998)  
 Other\_GSSs: CITBI-El-2517A7.TR  
 CONTACT: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: madams@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 End Search Page:  
[http://www.tigr.org/tcd/hunguen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tcd/hunguen/bac_end_search/bac_end_search.html).  
 Seq primer: M13-21  
 Class: BAC ends.  
 Location/Qualifiers  
 1..664  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="2517A7"  
 /sex="male"  
 /cell\_type="sperm"  
 /note="Vector: pBeloBAC11; Site\_1: KcoRI; Site\_2: RcoRI;  
 Caltech Human BAC Library D"

ORIGIN  
 Query Match 3.0%; Score 181.6; DB 28; Length 664;  
 Best Local Similarity 72.2%; Pred. No. 1.2e-16;  
 Matches 249; Conservative 0; Mismatches 95; Indels 1; Gaps 1;

QY 3283 TAGATGATGTGTGCTCATGCTGTAAATCCCAACCTTTGGGAAGCGGTCGGAGG 3342  
 DB 639 TGGCTGAGTGTGGGCTCACACCTGTAAATCCACACCTTTGGAGAGGCTCATGTGGAAG 580  
 QY 3343 ATAGCTGAGTCCAGAGCTTTGAGACAGTCAAGGCAACAGAGCAAGCCCATATCTAA 3402  
 DB 579 ACAGCTTAGGCCCAAGAGTTTGAATGACTAGCTTGGGCAACAGAGAAACCTCACTTAC 520  
 QY 3403 AAAACAAACAAACAAATTAATCTGGGATATGATGTGCTCACTGTAGTCAAGTAC 3462  
 DB 519 AAAAATAAATTTTAAAGAAAGTAAAGGAGTGTGAAGAAACCCATAGTCCAGTAC 460  
 QY 3463 ACAGAGCTGAGGCAAGAGATCACTTGAAGCTCCAGAGGTTTGAAGCTGAGTATTCAT 3522  
 DB 459 TCAGAGCTGAGGCAAGAGATCACTTGAAGCTCCAGAGGTTTGAAGCTGAGTATTCAT 400  
 QY 3523 GAACGCGTGTCACT-CAGTCTGGGTGACAGTCAAGAGCTGTCTCAAAAATATATATA 3581

```

Db      339  GATCATGCCACTGCTCCAGCTCGGSCACAGACGACGACCTGCTCAAAAAGAAAA 340
Qy      3582 ATAAATATAAATTAATTTTAAATAAATAAATAAATTAATTAATTTTA 3626
Db      339  GAAATATCAATTTGCTAGTTTCCCATCCTACCTACAGAAATTAAGTTTA 295

RESULT 8
CC061176
LOCUS    CC061176
DEFINITION MOCQ_CH252P002Q4T7_H05_CD314_036_CHORI-252 Vervet Monkey Library
          Cercopithecus aethiops genomic clone CH252-2B10, genomic survey
          sequence.
ACCESSION CC061176
VERSION   CC061176.1
KEYWORDS  GI:29790439
SOURCE    GSS.
ORGANISM  Cercopithecus aethiops (African green monkey)
          Cercopithecus aethiops
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
          Cercopithecinae; Cercopithecus.
REFERENCE 1 (bases 1 to 725)
          Minhas,R., Zhang,X., Dore,C., Villeneuve,A., Lepage,P.,
          Forgetta,V., McKee,K., Ophoff,R.A., Fairbanks,L.A., Freimer,N.B.,
          Ervin,F.R., Palmour,R.M., Hudson,T.J. and Dewar, K.
          UCDA/MOCQ/St-Kilts Vervet Monkey Mapping Project
          Unpublished (2003)
          Other GSSs: MOCQ_CH252P002Q4SP6_H05_CD315_036
          Contact: Dewar K
          McGill University and Genome Quebec Innovation Centre
          740 Dr. Penfield Room 7214, Montreal, QC, Canada, H3H 1M4
          Tel: 514 398 3311 x00089
          Fax: 514 398 1795
          Email: ken.dewar@mcgill.ca
          Plate: 2 row: B column: 10
          Seq primer: T7 : TAAATACGACTCATTATGCG
          Class: BAC ends.
FEATURES
          source
          Location/Qualifiers
            1..725
              /organism="Cercopithecus aethiops"
              /mol_type="genomic DNA"
              /db_xref="taxon:9534"
              /clone="CH252-2B10"
              /sex="male"
              /cell_type="white blood cell"
              /dev_stage="adult"
              /clone_lib="CHORI-252 Vervet Monkey Library"
              /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
              Constructed by Michael Nefedov in Pictet de Jong's
              laboratory at BACPAC Resources, Children's Hospital in
              Oakland Research Institute."
ORIGIN
Query Match      3.0%; Score 180.6; DB 26; Length 725;
Best Local Similarity 60.1%; Pred. No. 1.5e-16;
Matches 335; Conservative 0; Mismatches 219; Indels 3; Gaps 2;

Qy      3078 AACTCTGCCATTATTTCCATATGACATAGACCACTTTTCATTAATTAATTT 3137
Db      21  AACCCCTTAATTTATGCGACCTGTTATATTTCCATATATATTTTGTGTTGTT 80
Qy      3138 TAAATTTGTCGCAATTAATTTCTGTCCTCTAGCTACTTAATTAATTTTAA 3197
Db      81  TAAACAGCTATTAATTTTACAGATTAATTTTCTTTTGTCCTAGGTGAAGAACA 140
Qy      3198 TAAACAATTAATTAATGACATAGCTTAATTAATTAATTAATTAATTAATTAAT 3257
Db      14  TCTTCTGATTTTATGCTTTTAAATTTATTAATTAATTTTGTGACCTGGGTAAAA 200
Qy      3258 AAACCTTTTAAATTTTAAATAAATATGATGAGTGCTGCTGATCCCA 3317
Db      20  ACAATGAACATTAATTTATGATATATGCTGGGACAGATGACTGCTGTAATCCAG 260

```

```

Qy      3318 CACTTTGGGAAGCGCGGTCGGAGATAGCTTGAGTCCAGACTTTGAGACGACGACG 3377
Db      261  CAGTTTGGGAGGCGCAAGGGTGGATGATCACTTGAGATCCAGATTCAGACCGAGCTGGC 320
Qy      3378 CAACACGACGACGACCCCATATCTTAATAAACAACAAACAATAATTAATTAATTAAT 3437
Db      321  CAACATGGCAAAACCTCATCTCTAATAAATAAACAACAAACAAATAAATAAATAAATA 378
Qy      3438 TCTGCTGACCTGTTAGTCCAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 3497
Db      379  TGCACACACCTGTTAGTCCAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 438
Qy      3498 GGAAGTTGAGCTGCAAGTATCCATGAAGCGCTGCTGACT-CACTTGGGTGACAGTG 3556
Db      439  GGAGGACAGAGTTTACATGATGAGCAAGATTGCAACCATGCTCCAGCTTGGTGAACAG 498
Qy      3557 CAAGACCTGCTCAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3616
Db      499  AGAGACCTGCTCAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 558
Qy      3617 TTAATTTTAAACAC 3633
Db      559  TACAAATTTATGATAC 575

RESULT 9
BZ601705
LOCUS    BZ601705
DEFINITION MHABF61TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
          sapiens genomic clone MCF7_1-8K2, genomic survey sequence.
ACCESSION BZ601705
VERSION   BZ601705.1
KEYWORDS  GI:31510167
SOURCE    GSS.
ORGANISM  Homo sapiens (human)
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 828)
          Volk,J.S., Zhao,S., Chin,K., Brehner,J.H., Herndon,D.R., Tao,Q.,
          Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
          Gray,J.W. and Collins,C.
          End-sequence profiling: Sequence-based analysis of aberrant genomes
          Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7656-7701 (2003)
          12788976
          Contact: Volk J.S.
          Collin Collins' Lab
          UCSF Comprehensive Cancer Center
          UCSF Box 0808, San Francisco, CA 94143-0808, USA
          Tel: 415 502 7066
          Fax: 415 502 5665
          Email: svolik@cc.ucsf.edu
          This clone is available from Amplicon Express
          http://www.genomex.com
          Class: BAC ends.
FEATURES
          source
          Location/Qualifiers
            1..828
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /clone="MCF7_1-8K2"
              /sex="female"
              /clone_lib="Human MCF7 breast cancer cell line library
              (MCF7_1)"
              /note="Vector: pECHO1; Site 1: HindIII; This library was
              constructed from MCF7 breast cancer cell line by Amplicon
              Express (http://www.genomex.com) using their standard
              procedure."
ORIGIN
Query Match      3.0%; Score 180; DB 28; Length 828;
Best Local Similarity 67.7%; Pred. No. 1.7e-16;

```

Matches 283; Conservative 0; Mismatches 130; Indels 5; Gaps 2;

```

OY 3295 GTGGCTGATGCTGTATTCCTGATCCAGACCTTTGGGAGCGGGTCCGGAGAGTAGTGGATC 3354
    |||||
DB 401 GTGGCTGATGCTGTATTCCTGATCCAGACCTTTGGGAGCGGGTCCGGAGAGTAGTGGATC 460
    |||||
OY 3355 CAGCAGTTTGAGACCGTCAAGGCGACAGACAGACCCCAATCTTAAAAAACAACA 3414
    |||||
DB 461 CAGCAGTTTGAGACCGCTGGGCAACATGCTGACCCCATCT---ACATAAAAAAA 516
    |||||
OY 3415 AAACAAATATACCTGGGTATGGTGTGCTACCTGTATGCTCAAGCTACAGAGAGCTGA 3474
    |||||
DB 517 TAAATATGTTAGCTAGCTAGTATGATGACCTGTATGCTAGCTAGCTAGCTAGCTGA 576
    |||||
OY 3475 GCGAGAGATGACTGACCGGAGAGAGAGTGGAGCTGACATGATCCATGAAGCCCTGCT 3534
    |||||
DB 577 GCGAGAGATGACTGACCGGAGAGAGTGGAGCTGACATGATCCATGAAGCTGCACT 636
    |||||
OY 3535 ACATCT-AGTCTGGGTGACAGTGCAGAGAGCTGTCTCAAAATATATATATATATAT 3593
    |||||
DB 637 GCATCTAGCTGTGAGAGACAGAGAGAGTCTTGTCTCTTAAATATATATATATATAT 696
    |||||
OY 3594 AACTTAAAAAACAATATATATATATATATATATATATATATATATATATATATAT 3653
    |||||
DB 697 ATTTTAAAAAACAAGCTTGTGACTCTTGTGAGCAAGCTATATATATATATATATAT 756
    |||||
OY 3654 AATATGATTTTGGAGCTATATCTCTGAGAGTATATATATATATATATATATATAT 3711
    |||||
DB 757 AATCTTCTTTTGTGATGATATGATGATGATGATGATGATGATGATGATGATGATGAT 814
    |||||

```

## RESULT 10

LOCUS A0052879 733 bp DNA linear GSS 20-APR-1999  
 DEFINITION RPCI11-42023.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-42023,  
 genomic survey sequence.

ACCESSION A0052879  
 VERSION A0052879.1 GI:3349834

## KEYWORDS

## SOURCE

Homo sapiens (human)

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

ORIGIN /note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;  
 RPCI11 Human Male BAC Library"

Query Match 3.0%; Score 179.8; DB 28; Length 733;  
 Best Local Similarity 65.9%; Pred. No. 2e-16;  
 Matches 276; Conservative 0; Mismatches 142; Indels 1; Gaps 1;

```

OY 3196 AATAACAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3255
    |||||
DB 267 ATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 326
    |||||
OY 3256 ATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3315
    |||||
DB 327 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 386
    |||||
OY 3316 AACAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3375
    |||||
DB 387 AGCATTTGGAGGCGCCAGAGCTGAGAGACCTTGAAGCCAGAGTTGATGATGATGATGAT 446
    |||||
OY 3376 GCGACACAGACAGACCCCATATCTTAAAAAACAACAACAACAACAACAACAACAACA 3435
    |||||
DB 447 GACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 506
    |||||
OY 3436 GTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3495
    |||||
DB 507 GCAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 566
    |||||
OY 3496 CAGAGGTTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3554
    |||||
DB 567 CAGAGGTTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 626
    |||||

```

## RESULT 11

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

FEATURES  
 source  
 1..567  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6252453"  
 /tissue\_type="ductal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH\_MGC\_110"

High quality sequence stop: 542.  
 1..567  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6252453"  
 /tissue\_type="ductal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH\_MGC\_110"

BO691256 567 bp mRNA linear EST 15-JUL-2002  
 AGENCOURT 8176628 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6252453  
 5', mRNA sequence.  
 BO691256  
 BO691256.1 GI:21816572  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 567)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LCM2398 row: 1 column: 22





TITLE  
JOURNAL

Direct Submission  
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail:Chimpbes@sec.riken.go.jp, URL:http://hsp.sec.riken.go.jp/,  
Tel:81-45-503-9111, Fax:81-45-503-9170)  
Clones are derived from the chimpanzee BAC library PTB This BAC end  
was generated during the R&D process and may have higher chance of  
clone tracking errors.

## COMMENT

## PRIMERS

Sequencing: -21M13

## LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1..674

/organism="Pan troglodytes"

/mol type="genomic DNA"

/db xref="taxon:9598"

/clone="PTB-065D24.F"

/sex="male"

/cell\_type="lymphoblast"

/clone\_id="PTB Chimpanzee Male BAC library"

## ORIGIN

## Query Match

3.0%; Score 178.2; DB 29; Length 674;

Best Local Similarity 75.9%; Pred. No. 3.6e-16;

Matches 233; Conservative 0; Mismatches 73; Indels 1; Gaps 1;

```

QY 3289 AGTGTGGTGGCTCATGCTGTATATCCCAACACTTTGGAGACCGGGTGGAGGATAGCT 3348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 336 AGTGAAGTGGCTCATGCTGTATATCCCAACACTTTGGAGACCGGGAGGATGACT 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3349 TGAATCCAGCGAGTTTGAGACCACTGAGGCAACAGCAAGCCCATATCTAATAAAC 3408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 336 TGAATCCAGCGAGTTTGAGACCACTGAGGCAACAGCAAGCCCATATCTAATAAAC 278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3409 AAAACAAACAAATTAACCTGGGTATGTGTGCTACCTGTAGTCCAGCTACACAGGA 3468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 277 ATAAATAATTAATAATGAGCCAGAGTGTGTGCTACCTGTAGTCCAGCTACACAGGA 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3469 AGCTAGGCGAAGAGATCACTTGAGCCGAGAGGTTGAGGCTGCAAGTATCCATGAACGC 3528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 217 GCTGAGGTGGAGATCACTTGAGCCGAGAGGTTGAGGCTGCAAGTATCCATGAAGTAT 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3529 GCTGTAAGTCACTGCTGGGTGACAGTGCAGAGAGCTGTCTAATAATAATAATAATA 3588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 157 GCGACTGCACTCAATCTGGGTGATAGAGTGAAGCCCTGTCTCAAAAAAATTAATAAGA 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3589 AAAATAA 3595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 97 ACTATGA 91
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: September 27, 2004, 18:27:21  
Job time : 13719 secs